REPORT DOCUMENTATION PAGE AFRL-SR-AR-TR-06-0083

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, sea ... eviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including sugg ... ormation Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Pape.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 2006	FINAL

4. TITLE AND SUBTITLE
Molecular Mechanisms and Modeling of Skin Irritation from JP-8

5. FUNDING NUMBERS
F49620-03-1-0115

6. AUTHOR(S)
James N. McDougal

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Pharmacology
Boonshoft School of Medicine
Wright State University
3640 Col Glenn Hwy
Dayton, OH 45435

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NL
875 N. Randolph Street
Arlington, VA 22203
Dr. Walter Kozumbo

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approve for Public Release: Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

AFOSR Final progress Report

Molecular Mechanisms and Modeling of Skin Irritation from JP-8

F49620-03-1-0115

James N. McDougal, PhD,
Department of Pharmacology,
Boonshoft School of Medicine
Wright State University,
3640 Col Glenn Hwy,
Dayton OH 45435

Period covered: **February 28, 2003 to February 9, 2006**

This project investigated many aspects of JP-8 induced skin irritation at a molecular level. Studies focused primarily on the effects of short-term *in vivo* exposures in rats. Exposure methods were developed and refined to assure that the animals could be exposed cutaneously in a repeatable manner while they were in their home cages. Methods were developed to separate the epidermis from the dermis so that the molecular responses could be determined separately. Gene expression (Gene chips and RT-PCR) and changes in protein levels (ELISA) due to exposures to JP-8 and some of the primary components were identified. These results showed that the response of the skin to a one-hour exposure to JP-8 can be quantified and increases for up to eight hours after exposure. The specific genes involved change with time. Early responses primarily involve transcription factors and DNA binding genes and the later responses involve transcripts related to metabolism and oxidative stress. Levels of JP-8 and identifiable components in the epidermis and dermis were quantified by GC mass spectrometry. During and after the JP-8 exposure aliphatic components are found in the skin at several times the concentration of the aromatic components of JP-8. Individual gene expression studies with two aliphatic and two aromatic JP-8 components suggested that chemical characteristics such as leaving groups and a benzene ring were more important in triggering the inflammatory response that the general category of chemical. The suspected trigger of the inflammatory process (IL-1α) was localized to the stratum corneum and viable epidermis with immunohisto-chemistry. A systems biology model for the primary irritant pathway in the skin was developed in coordination with the CIIT Centers for Health Research. Sensitivity analysis of this model identified parameters that should be refined with laboratory studies. Preliminary approaches to manipulate the irritant pathway in the skin using intradermal injections were developed.

Published manuscripts:

Rogers, J.V. and J.N. McDougal. Improved Method for in vitro Assessment of Dermal Toxicity for Volatile Organic Chemicals. Toxicology Letters, 135:125-135 (2002).

20060405013

Cover and title page SF 298 Report Document Page

100-200 word abstract for this page

This report describes a program of research investigating collaborative memory processes. The program's aim was to understand how collaboration among members of AWACS weapons director teams could influence their performance. Predictions for team memory performance were derived from social decision scheme theory and the ideal group model derived from signal detection theory. These theoretical approaches were used to establish benchmarks for optimal, suboptimal, and lower-bound levels of team performance. The research demonstrates that team performance exceeded that of similarly treated individuals. Moreover, some results suggest that teams might exceed the predictions of optimal performance from the ideal group model. Results from this research also indicated that larger teams achieve better performance on various metrics than teams with fewer members. Additionally, results are reported describing how the characteristics of the information displayed influenced memory rates. In particular, information about an aircraft's mission, location, and identification were remembered much better than numeric information about heading and altitude. This research effort demonstrated these findings in a synthetic task environment that resembles the cognitive characteristics faced by AWACS weapons directors, particularly requirements for memory performance. This research facilitates the development of knowledge of how collaboration can be used to harness the advantages of teams.

2. Executive Summary

Final Report: Modeling Memory Processes and Performance Benchmarks of AWACS Weapons Director Teams

Memory is a critical cognitive process in the multiple tasks performed by Air Force AWACS weapons directors. When weapons directors lose access to the screen display information, they must rely upon their memory to perform their tasks. Weapons directors can respond to this situation by using collaborative strategies to enhance their team's memory performance. This report describes a program of research investigating the impact of collaboration on memory performance. Among the unique features of this research is the development and test of benchmarks for team performance based on theoretical models derived from signal detection theory and social decision schemes theory. These benchmarks defined optimal and suboptimal levels of performance on the memory tasks. Across the studies reported, optimal performance was found for six member groups responding to information from a mock job interview while suboptimal performance was found for three- and five-person teams answering questions about difficult to remember information displayed in an AWACS weapons directors synthetic task environment. These patterns of results suggest that theoretically-derived benchmarks provide useful means for defining optimal and suboptimal levels of performance. Moreover, the comparison of performance to the benchmarks helps highlight and identify conditions under which teams do and do not approximate optimal performance. A direction for new research is to capture the expertise of weapons directors to enhance team performance.

Results of the studies reported indicate that the nature of the information to be remembered as well as team size interacted to influence the social and cognitive processes by which teams responded to memory items. For the studies using rich, but not extremely challenging material, the group responses were influenced by members being able to know the correct response and demonstrate it to other group members. For the study involving the very challenging information displayed to simulated weapons directors, the team responded according to a majority process regardless of whether the majority was correct or incorrect. The results of these studies have an additional implication for considerations of reducing weapons director team size. That is, weapons directors face such a cognitively demanding set of tasks, reducing team size would diminish their team effectiveness. The theoretical models considered and the empirical evidence gathered both support keeping the proper team size for weapons directors unless substantial changes are made in terms of memory aids and the way information is displayed.

Additional results of these studies reinforce the finding that the nature of the information displayed influenced memory performance. Based on theoretical conceptualizations of the types of information displayed to weapons directors, it was predicted that certain attributes (e.g., aircraft mission) would be more likely to be remembered than other information (e.g., aircraft heading). The results of one study demonstrated strong support for this set of predictions. Even when the simulated weapons directors were faced with strong cognitive demands, certain aircraft attributes were significantly more likely to be remembered correctly than information about other attributes. These findings suggest that specific modifications of the way information is displayed to weapons directors could aid them in being able to remember the information and perform their tasks more effectively. In addition to these conclusions about display information and team size, the studies reported also contribute to our theoretical understanding of team performance and help build a stronger conceptualization of how collaboration can be used to harness the advantages of teams.

3. **Table of Contents**

		Page Number
1.	Report Documentation Page (SF 298)	1
2.	Executive Summary	2
3.	Table of Contents	3
4.	Background	4
5.	Objectives	4
6.	Overview of Research Effort	5
7.	The Ideal Group Model Analysis and Test	7
8.	Further Investigation of the Ideal Group Model	13
9.	Social Decision Schemes and Team Decision Making	21
10.	Team Performance among AWACS Weapons Directors	23
11.	Differential Memory for Weapons Directors Display Information	39
12.	Summary and Conclusions	46
13.	Supporting Efforts in Group Decision Making	49
14.	Supporting Efforts in Group and Team Performance	49
15.	Supporting Efforts in Group and Individual Goal Setting	50
16.	Supporting Efforts in other Judgment and Decision Making topics	51
17.	Additional Efforts in other Social-Organizational Psychology topics	51
18.	Development of Students	52
19.	Presentations Made During Grant Funded Period	53
20.	Literature Cited	59
21.	Personnel Supported by Grant Funds	63
22.	Appendix A	64
23.	Appendix B	68

4. Background

The proposal that led to this funding was submitted in July 2001 by Verlin Hinsz as Principal Investigator. The proposal was inspired by presentations and discussions about the AFOSR Team Performance program held at Brooks AFB in April 2001. Verlin Hinsz began a 12 month NRC/AFRL Senior Research Associateship in August 2001. Dr. Hinsz held this position at the AFRL/HEA Team Training branch at Brooks AFB with Sam Schiflett as his advisor. The time Dr. Hinsz spent as a Senior Research Associate was very informative and aided in the work conducted as part of this funding.

The submitted proposal was selected for 36 months of funding for the period beginning June 2002. A request for an extension was made in early 2005 to help finalize the project and account for the funding. The extension was requested in part because North Dakota State University (NDSU) was undergoing a change in accounting systems to PeopleSoft and it was not clear what the balance in the grant was. Consequently, the funding period for this grant ended on the revised date of October 31, 2005.

When this proposal was funded, Dr. Hinsz held an NRC Senior Research Associateship. One condition of this Associateship was that he could not receive other federal funds at the same time. Consequently, although conceptual thought and effort could be dedicated to the project, no expenditures could be made. Dr. Hinsz also spent some time trying to recruit student and technical personnel to work on this project during the initial months. Once the Associateship ended in August 2002, Dr. Hinsz returned to NDSU and began working on the project in earnest.

The research in the project was conducted at NDSU. Dr. Hinsz held a position as professor of psychology in the College of Science and Mathematics at NDSU. Space for this research effort was made available in Hultz Hall on the NDSU campus. In this space, the two rooms were arranged to make an AWACS Team Performance Laboratory. A set of workstations were erected in each of these rooms to test team performance in a simulated AWACS Weapons Director team environment. More details of this lab are presented later in this report (Section 10).

The simulated AWACS Weapons Direct synthetic task environment used for some of this research was produced under a subcontract with Aptima Incorporated. The weapons director synthetic task environment (STE) used in this research effort was a variation on the DDD AWACS Weapons Director STE that had been developed for AFOSR research (cf., Entin, Entin, Bailey, Paley, & Miller, 2001; Entin & Rubineau, 2002). It took officials at NDSU four months to approve the subcontract to Aptima to construct a revised DDD AWACS STE. This local delay was compounded by repeated delays in Aptima producing the revised DDD STE program. Although the program should have been completed in several months, it appeared that Aptima personnel were directed toward other efforts while the subcontract was in effect. Consequently, NDSU did not receive completely functioning DDD AWACS scenarios until approximately March 2004. Because of these long delays, the personnel supported by this funding completed research related to the primary objectives so that important conceptual questions could be addressed. The objectives and the research efforts pursued are described next in this report.

5. Objectives

The reported research investigates collaborative memory processes associated with the performance of weapons director teams. Predictions for team memory performance are derived from the combinations-of-contributions theory model of team performance (Hinsz, Tindale, & Vollrath, 1997). Comparisons between individual and team performance as well as differences between teams of different sizes are described. In addition, models of memory performance in

teams are specified based on the ideal group model (Sorkin & Dai, 1994) and social decision schemes theory (Davis, 1973; Hinsz, 1990). These models establish benchmarks for optimal, sub-optimal, and lower-bound levels of performance. The experiments explore the degree the performance of weapons director teams approaches the various benchmarks for performance.

The experiments also investigate the possibility of performance that exceeds optimal levels as defined by the models. The research could then investigate the processes that may contribute to this extraordinary performance in weapons director teams. Moreover, examination of the optimal and baseline levels of memory performance can help identify why teams may not attain optimal levels of performance. The models applied in this research also predict the impact of reduced team size on the memory performance of AWACS weapons director teams.

The research reported also investigates a number of features of memory processes associated with performance in weapons director teams: errors of omission, errors of commission, sensitivity, and bias. The responses to memory items also provide an opportunity to assess different types of memory that might be involved in weapons directors' task. The information displayed to weapons directors can be symbolic, semantic, numeric, spatial, and alphanumeric. The responses provide evidence regarding differences in memory for these aspects of the weapons directors' tasks.

This research helps advance the theoretical understanding of team performance and allows consideration of the efficacy of different formal models of team performance. Some of this research relies upon a synthetic task environment to allow comparisons with the results of other research. This research facilitates the development of knowledge of how memory and collaboration in memory can be used to harness the advantages of teams.

6. Overview of Research Effort

Weapons directors are Air Force personnel that control and direct aircraft assets (e.g., fighters, refueling tankers) from airborne platforms (AWACS: Airborne Warning and Control Stations). The tasks weapons directors perform are very complex, involving an array of cognitive abilities (Fahey, Rowe, Dunlap, & deBoom, 2000). Weapons directors also collaborate and coordinate with other weapons directors, the senior director, and personnel operating other aircraft to insure the safety of the aircraft and personnel. Weapons directors sit at workstations and monitor the activities of aircraft that are their responsibility from screen displays. With adequate training and proper design of instruments, weapons directors can perform their tasks quite well. However, at particular times, weapons directors can lose their screen displays and have to perform their tasks "blind". In these conditions, weapons directors need to rely on their memory to be able to monitor, command, control, and communicate with the personnel on other aircraft. In this critical situation, the members of a weapons director team may be able to use other team members to aid them in remembering where aircraft are, what their missions are, where they are heading, possible conflicts that might exist in the flight corridor, and the existence of threatening conditions (e.g., fighter running out of fuel). This research effort aimed to investigate how collaboration in memory performance of weapons directors enhances their performance.

Memory is a critical cognitive process in the multiple tasks performed by weapons directors (Fahey, et al., 2000). Memory processes are critical for at least 10 different areas of the task performance of weapons directors in AWACS environments (Hinsz & Malone, 2004). One way to improve the memory processes of weapons directors is to take advantage of their team structure and use collaboration to enhance memory performance. Research on teams, groups, and crews provides a strong theoretical and conceptual basis for the consideration of the memory

performance of weapons director teams. Signal detection theory can be applied to memory performance in teams (Banks, 1970; Green & Swets, 1966; Hinsz, 1990). Moreover, social decision scheme theory relates to team judgments about recognition items (Davis, 1973; Hinsz, 1990, 1999). Much of this research effort follows from the combinations-of-contributions meta-theory for team processes and performance (Hinsz, Tindale, & Vollrath, 1997; Hinsz, Vollrath, Nagao, & Davis, 1988).

The theory of combinations of contributions argues that team performance on a task can be understood and predicted based on two components. The first component is the <u>contributions</u> that members bring to the team. In the case of team memory for information displayed on weapons directors' screens, the weapons directors' contributions would be their memories of the displayed information. The second component is the ways these contributions are <u>combined</u>. Combinations-of-contributions theory argues that the way the contributions are combined can be described in some systematic fashion as a function of the task involved (e.g., average or sum of contributions). According to combinations-of-contributions theory, the important questions to be addressed in weapons directors' memory performance are: (a) what do the weapons directors remember about the displayed information, and (b) how do they combine the information to produce accurate or inaccurate memory responses.

Research examining team performance has often considered the comparison of the relative performance of teams and individuals (see Guzzo & Salas, 1995; Ilgen, 1999). If teams are asked to perform a task, it is important to know whether teams achieve more than individuals that are treated similarly (cf., Davis, 1969). If teams fail to outperform individuals, then one must question whether it is appropriate to have teams work at the task (Hackman, 1987, 1990). Fortunately, research results generally indicate that teams are superior to individuals on performance of cognitive tasks (Hill, 1982; Hinsz, et al., 1997; McGrath, 1984). Consequently, research has attempted to identify the mechanisms and processes that lead to this superiority of team performance.

Hinsz (1990) suggested three bases for the superiority of teams in performance on cognitive tasks: *information pooling*, *error correction*, and *effective strategies*. Teams can be superior to individuals because they have a larger information pool upon which to draw when performing a cognitive task. Teams also improve upon individuals by correcting errors in their members' responses. Because team members can identify the errors of fellow members, they are less likely to incorporate them in the team response. Teams also appear to use better and more effective strategies for selecting an appropriate response than do individuals. These three mechanisms can work independently and interactively to enhance team performance on cognitive tasks such that it is superior to that of similarly treated individuals. The experiments described in this report examine how information pooling, error correction, and effective strategies contribute to the memory performance of weapons director teams.

The members of weapons director teams can be treated as individuals or as a team in terms of their memory for the information displayed at their workstations. As weapons directors control and monitor aircraft, the information they are presented is basically the same. Consequently, in terms of combinations-of-contributions theory, individuals and team members would have virtually equivalent information that could serve as contributions to individual or team performance on the task. Therefore, according to combinations-of-contributions theory, the ways in which team and individual performance differs is in terms of how the information from the team members is combined. The theoretical models applied here focus on how weapons directors' contributions are combined and how bringing weapons directors together as a team may transform these contributions.

The theoretical approaches considered in this proposal have both prescriptive and

descriptive aspects. One aim of this research effort is to establish benchmarks that can be used to evaluate the performance of weapons director teams. A number of different benchmarks will be considered: average performance under standard conditions, optimal levels of performance, and lower-bounds for expected levels of performance. In addition, the models allow us to consider what would be required to attain levels of performance that exceeds optimal according to the models.

The theoretical models applied to the responses of weapons directors provide a means for generating benchmarks for optimal levels of performance. These benchmarks for optimal performance describe what would be required for the teams to achieve maximum efficiency and optimum use of member contributions. To improve the performance of weapons director teams, it is beneficial to know the degree to which interventions lead teams to approach optimal levels of performance. Both signal detection theory and social decision scheme theory provide conceptual bases upon which to specify optimal levels of performance given the weapons directors' individual performance levels.

The theoretical approaches applied in these experiments also provide a basis for considering non-optimal levels of performance. It might be important to recognize how poorly a team performs relative to sub-optimal benchmarks for performance. The models considered can define lower bounds of performance that could be expected if specific team processes occurred in weapons director teams. Consequently, it might be possible to identify processes of team memory performance that are leading the weapons directors to fail in their task performance.

An objective of this research effort is to describe the basic features of the memory performance of weapons director teams. This includes the average level of memory performance (e.g., proportion correct) and other indices used to reflect aspects of the teams' memory processes (e.g., d'). These experiments will also indicate the frequency of different types of errors (e.g., errors of omission and commission) in the memory responses as well as bias in types of errors (e.g., $\exists$). Thus, a number of features of the memory processes of weapons director teams are revealed as a function of the research reported.

7. <u>The Ideal Group Model Test and Analysis</u>

One approach used for considering memory performance in weapons director teams is based on signal detection theory (see Hinsz, 1990; Sorkin & Dai, 1994; Sorkin, Hays, & West, 2001). The application of signal detection theory in this research is consistent with the combinations-of-contributions formulation. The contributions of team members are the sensitivity and accuracy of the judgments they make for the memory responses. The combination aspect of signal detection theory reflects the models that integrate the responses or information available to the team members (Sorkin, West, & Robinson, 1998). The integration model of signal detection theory also represents the information pooling thought to contribute to the superiority of team performance.

Signal detection theory allows investigation of the degree individuals and teams compare in terms of their sensitivity to the information displayed to weapons directors. The index of sensitivity, d', reflects the degree that team members correctly recognize the information as displayed. Signal detection theory also provides an index, $\exists$, that reflects the degree and types of bias in the memory responses (i.e., the relative proportion of miss and false alarm responses). The d' and $\exists$ indices provide ways to compare how teams and individuals respond to recognition memory items. Given the expected superiority of team performance, the team d' values should be greater than those of the individuals.

Another distinct advantage of signal detection theory is that an extension of the multiple

channel model can be applied to define theoretically optimal levels of performance (*d'*) that would be expected from a team if it optimally combined and used the available information (Green & Swets, 1966; Sorkin & Dai, 1994; Sorkin, Hayes, & West, 2001). Sorkin and Dai (1994) argue that this model is applicable to a variety of teams, including flight crews. Consequently, this signal detection approach appears suitable to consider team responses of weapons directors responding to recognition memory items.

One of the achievements of Sorkin's analysis is that it defines the ideal (optimal) *d'* for a team (the ideal group model). The ideal team *d'* is calculated based on four parameters: The team size (*m*), the mean detectability of the team members ($\bar{d}'$), the variance of the member detectability indices (var *d'*), and the intercorrelation of the responses of the team members (*r*). The intercorrelation of the *m* team members would be the intraclass correlation (ICC) calculated based on the *m* team member responses to the set of items used to determine the detectability index, *d'* (see Shrout & Fleiss, 1979, for descriptions of intraclass correlations). The model is expressed as:

$$\text{ideal group } d' = \sqrt{\left[\frac{\mathrm{m}\left(\mathrm{var}\ \mathrm{d}'\right)}{1-\mathrm{r}} + \frac{\mathrm{m}\left(\bar{\mathrm{d}}'\right)^2}{1+\mathrm{r}(\mathrm{m}-1)}\right]} \qquad \text{(Eq 1.)}$$

From this equation, it can be discerned that ideal team *d'* increases as team size, mean *d'*, and var(*d'*) increase, and as *r* decreases.

The ideal team model predicts the potential level of performance if all the information available among the team members is applied accurately for the group response and if the team members' contributions are weighted to produce the optimal combination of member knowledge. This ideal team model is remarkable in that it is also a model of the four variables that are important for predicting team performance in a signal detection study. As such, the ideal team model is a model of team performance on cognitive or perceptual tasks as well as a model that defines ideal team performance in such studies.

Sorkin and his colleagues also propose a way of evaluating the efficiency of team performance relative to that expected if the team operated ideally. Sorkin and Dai (1994, Eq. 20) offer an efficiency index, η, which is the square of the observed team *d'* divided by the square of the ideal team *d'*. This team efficiency index provides another way in which the observed performance of teams can be compared to predictions of an ideal team to gain a better understanding of how actual teams may fail to operate at ideal levels.

The ideal team *d'* model provides a basis for analyzing data reported earlier (Hinsz, 1990) to determine how well it accounts for the observed group performance on a recognition memory task. This analysis can also reveal how well the observed group performance approximates potential group performance as defined by the ideal group model. Consequently, next the relevant aspects of the Hinsz (1990) study design and methods are briefly summarized. Additionally, some of the basic findings of the study are provided to put the signal detection analyses in context. More details regarding the design, method, and findings of the study can be found in Hinsz (1990) which also gives greater consideration to group memory performance.

Method of Hinsz (1990)

Student volunteers met in a large room and were asked to sit around a large table. They were provided introductory instructions and then asked to watch an 18 minute recording of a job interview on a video monitor. The participants were instructed that they would be tested on their

memory of the content of the video recording. Once the video finished, the students were randomly assigned to rooms with six participants in each room for an eventual group (55 groups, $n = 330$) or at least two students in a room for the individual condition ($n = 60$). Once in the room, each student received the individual recognition memory test and a response form. The recognition tests consisted of 60 items varying in difficulty based on statements and observations from the video recording. The students responded to the recognition items on computer-readable forms indicating whether they thought the item was true or false. The students were given extensive written and oral instructions for making their true/false judgments.

After completing the initial recognition test, the students in the group condition received a very similar second test and responded to this second test as a group. The group members were told to read each item together and to discuss the item among them to come to a group consensus regarding whether the item was true or false. Students in the individual condition completed the second recognition test again as individuals without interaction. At the conclusion of the session, the participants were debriefed and thanked for their participation as they were excused from the experimental session.

Results

Group versus Individual Memory Performance

Analyses indicated that individuals (69% correct) and eventual group members (68% correct) were somewhat accurate on their responses to the first recognition test, $t\ (338) = 0.25, p > .80$. However, groups were more often correct on the second test (84% correct) than the individuals (68% correct), $t\ (113) = 13.74, p < .0001$. The individuals' (mean $d' = 1.03$) and group members' (mean $d' = 1.04$) initial responses to the recognition test indicated that they were sensitive to the material presented, $t\ (387) = 0.20, p > .84$. The individuals' sensitivity to the material did not change over time (second test $d' = 1.02$). However, groups had significantly better memory performance (mean $d' = 2.11$) than the individuals, $t\ (113) = 13.29, p < .0001$. The β measure of bias in responding true or false did not differ between group members and individuals on the first test (0.87 vs. 0.86), although the groups and individuals became significantly more biased to respond with false alarms (i.e., true) on the second test (0.79 vs. 0.81), $F\ (1, 112) = 7.99, p < .01$.

Ideal Group Analysis

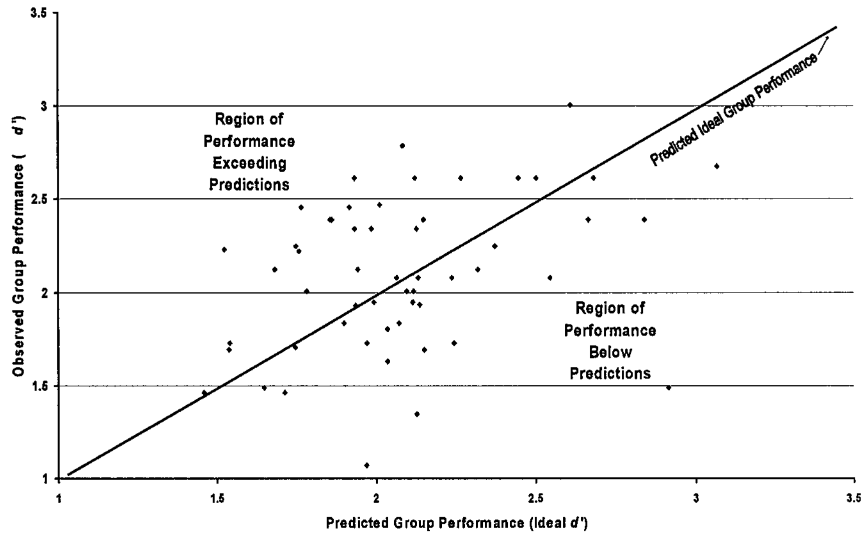
The calculation of the ideal group d' for this sample of groups was based on group size of $m = 6$, intraclass correlation among the group members' initial test responses = .25, and the mean group member $d' = 1.04$ with a variance in these group member $d' = 0.19$. Applying these values revealed an ideal group $d' = 2.11$, which is equal to the mean observed group $d' = 2.11$. This is a very surprising result suggesting that groups responding to the recognition memory test were integrating their information at levels predicted by the ideal group model. Moreover, this result suggests that the ideal group model is an exceptional model of group recognition memory performance for this sample of groups.

Group efficiency in integrating the information can be calculated from the ideal and observed group d' values. Not surprisingly, the groups achieved 100% of efficiency. This finding runs counter to most existing literature by suggesting that these groups achieved their potential levels of performance. That is, these groups did not appear to suffer from coordination, process, motivation, or other losses (Steiner, 1972) when integrating the information from their members' memory responses.

Given the equivalence of the mean observed group d' and the ideal group d' value in this analysis based on the total sample, it is of interest to determine the distribution of the 55 observed group d' values and their associated ideal group d' values. Because the overall observed

and ideal group *d'* values were equal, two possible distribution patterns might be anticipated. In the first pattern, every group would have achieved levels of performance predicted by the ideal group model in its responses to the recognition memory test. The second pattern suggests that the observed group *d'* are distributed around the mean, with perhaps half of the groups achieving performance at levels greater than those predicted by the ideal group model.

The following analysis describes the distribution of the observed group *d'* and related values relative to ideal group *d'* and for the 55 groups. Figure 1 presents the observed group *d'* values relative to those predicted by the ideal group model for the parameters given for that group. Note that the diagonal indicates a line of predicted ideal group performance if group performance achieved that of the ideal. Values to the left of that line indicate groups whose performance exceeded that predicted by the ideal group model. The region to the right of the diagonal line indicates suboptimal group performance relative to that predicted by the ideal group model.



The observed group *d'* values ranged from 1.07 to 3.00 while the ideal group *d'* for the 55 groups ranged from 1.46 to 3.07 with a mean = 2.08. The ideal group *d'* values were based on intraclass correlations with a mean of .26 and a range of .03 to .41, mean member *d'* of 1.04 that ranged from 0.50 to 1.47, and variance in group member *d'* of .17 that ranged from .03 to 1.16 for the 55 groups. When the group efficiency indices were calculated for the 55 groups, the mean efficiency was 111% and ranged from 26% to 214% efficient. For 28 of the 55 groups, the observed *d'* exceeded their predicted ideal group *d'*. These results are very intriguing and perplexing because they suggest that about half of the groups performed at levels exceeding those predicted by the ideal group model.

Discussion

The results of these analyses suggest that groups can perform at theoretically ideal levels on a recognition memory task. These analyses also suggest that the ideal group model does exceptionally well in accounting for information integration in group recognition memory performance. Additionally, under the assumptions of the ideal group *d'* model, these data suggest groups may perform at levels that exceed expectations. The findings of group performance meeting and exceeding predicted ideal levels are both radical departures from the traditional findings and conceptions of group performance as suboptimal (Steiner, 1972).

The ideal group model is a compelling approach for the consideration of group performance. It accounts for observed patterns of group performance and predicts ideal levels of performance for these groups. The ideal group model also incorporates the intercorrelation among group member responses along with group size, member competence, and heterogeneity of member contributions in an elegant theoretical formulation. This ideal group model proposes how each factor alone, and in combination, influences group performance (Sorkin et al., 2000). It

is not surprising that the competence of the group members on the task, d', contributes to overall group performance. Member ability has long been considered to be a critical factor in group performance (Davis, 1969). Moreover, the impact of variance in member d' values is anticipated by research that finds heterogeneity in member knowledge contributes more to proper solutions than does homogeneity (Shaw, 1981). Increases in group size have also been predicted and found to enhance group performance (Davis, 1982). The intercorrelation of eventual group member responses has not attracted as much attention in the prediction of group performance, but Sorkin's research demonstrates its impact.

The feat of groups performing at levels predicted as ideal is surprising given historical findings that group performance is suboptimal and often worse than the performance of the best member (Davis, 1969, 1982; Hastie, 1986; Hill, 1982; Steiner, 1972). These results are even more surprising given that these groups were randomly assembled ad-hoc groups of undergraduate students with no history or expectations of future interaction. Clearly, aspects of group recognition memory and the ideal group d' model deserve further examination to determine how the predicted ideal performance was achieved. If future research can identify the mechanisms leading to this ideal performance and interventions can be constructed to harness them, innovative approaches to the application of groups and teams will result. However, first this exceptional group performance needs to be replicated and understood.

The empirical demonstration of groups performing beyond their assumed potential is striking and defies comprehension within current conceptions of group performance. For example, the combinations-of-contributions approach implies that group performance is a combination of the contributions that group members bring to the group. The ideal group model, as one example of a combination-of-contributions approach, does not allow for performance beyond ideal levels (Sorkin et al., 2000). If the exceptional group performance identified here is defended with additional research, then it would violate the theoretical premises of the combinations-of-contributions approach.

Potential Explanations for the Exceptional Group Performance

The group performance exceeding predictions of the ideal group model reported above cannot be simply dismissed. It deserves explanation. The group performance beyond ideal levels observed with these data may result from violations of the assumptions of the theory of signal detection. However, further inspection of potential violations of the assumptions suggests they are not credible or imply that even higher group efficiency indices would have been found. Similarly, the group performance beyond the predicted ideal might result from violations of underlying assumptions of the ideal group model. Again, further analyses indicate that potential violations of assumptions would have had minor influence on the values calculated for the ideal group d'. Additionally, unreliable estimates of the critical parameters of the ideal group model appear unlikely given the sample of 55 groups. Consequently, the current focus will turn to the cognitive processes and motivational influences that emerge in interacting groups as more psychologically interesting ways of attempting to explain the observed group performance beyond theoretical expectations.

Stimulated Cognition in Group Memory Performance. Hill (1982) proposed that group members can stimulate the cognitive processes of other group members to benefit their group's performance. Groups may achieve performance exceeding predictions of the ideal group model because they could engage in a process similar to cued recall (Hinsz et al., 1997; Meudell, Hitch, & Boyle, 1995). Memory researchers have considered the possibility that collaboration in remembering can produce socially-cued memory. For example, when a group confronts a recognition item, one member might remember one piece of information relevant for answering the question. That member reports his or her memory to the other group members which could

cue or stimulate the other group members to remember material they were unable to retrieve earlier. In this fashion, group members may not have initially responded to the item correctly as individuals, but would respond correctly as a group. If this socially-stimulated cognition occurred for a number of recognition items, then group performance could exceed that predicted by the ideal group model. This socially-stimulated cognition explanation suggests that performance beyond that predicted by the ideal group model could arise because groups use this type of effective strategy in their responses.

Although there is a conceptual basis for the influence of socially-stimulated cognition, research has not revealed much support for it (Hinsz et al., 1997; Meudell, Hitch, & Boyle, 1995). Research results more generally find that group interaction hinders recall (Weldon, Blair, & Huebsch, 2000). The retrieval and announcement of one piece of information by a group member may interfere with another member's memory performance (Basden, Basden, & Henry, 2000). Although group memory performance exceeding the predicted ideal could result from socially-stimulated cognition, it is more likely that group interaction would interfere with the memory performance of group members (e.g., Basden, et al., 2000). Nevertheless, one worthy direction for future investigations of group memory performance is the potential impact of socially-stimulated cognition which could contribute to exceptional performance.

Motivational Influences of Group Interaction. One plausible way in which group performance could exceed ideal predictions is to assume that asking the individuals to make a group judgment motivates their memory for the material. This explanation suggests that forming a group of individuals increases the group members' motivation to remember the material when they respond as a group. If this motivational enhancement occurred in the groups, then the group members' initial d' values would underestimate the members' actual memory for the information. Consequently, the predicted ideal group d' would be an underestimate of potential group performance on the second memory test.

There are a few instances in which asking individuals to respond as a group appears to raise the motivation of individual members (i.e., social compensation, Karau & Williams, 1997; Köehler effect, Hertel, Kerr, & Messé, 2000; group brainstorming, Paulus, Dugosh, Dzindolet, Coskun, & Putman, 2002). Although it has been considered that groups can motivate their members (Hackman, 1992), research generally finds that groups reduce the motivation of individual members for task performance (e.g., social loafing, Latané, Williams, & Harkins, 1979). Sorkin et al. (2000) argue that a portion of the suboptimal performance they observed was due to motivation losses that led their observers to put forth less effort on some trials. Moreover, collaborative memory research has found that group performance can be less than baselines from individual memory performance (Clark, Hori, Putnam, & Martin, 2000; Weldon, Blair, & Huebsch, 2000). Consequently, a plausible explanation of beyond ideal group performance is that having the participants respond as a group may have motivated them to do well in their memory performance. However, there is substantial research suggesting that the opposite effect is more likely. Nevertheless, the entry of individuals into groups can have dramatic motivational influences for the way groups respond to memory questions and deserves further investigation. The following study examines this potential for enhanced motivation in groups to influence group recognition memory performance.

Summary

The findings reported here expand upon earlier findings that groups were superior to individuals in memory performance (Clark & Stephenson, 1989; Hinsz, 1990; Vollrath et al., 1989). As a model of the aggregate data, the predictions from ideal group model had an amazingly good fit. These groups' performance at levels that met and exceeded those of the predicted ideal is inconsistent with the combination-of-contributions metatheoretical approach,

and challenge longstanding conceptions of the nature of group performance. Potential explanations for the exceptional group performance reported based on potential motivational influences or socially-cued cognition appear to provide the best rationales for the exceptional group performance. These issues need to be resolved with future research, but this study does provide some bases for considering the potential of exceptional group performance on cognitive tasks.

8. Further Investigation of the Ideal Group Model

The discussion directly above highlighted a study that showed groups performed at extraordinary levels relative to predictions of the ideal group model. Two issues relating to that study were discussed that lead to the current investigation. One issue is that the finding of group performance that met and exceeded the performance predicted by the ideal group model should be replicated. A second issue to be addressed is whether the extraordinary performance might be explained by motivational influences on group members during interaction. The replication and potential explanation of the extraordinary performance will be considered in this study along with related issues. [This discussion is based upon a study partially reported in Hinsz, Spieker, Heimerdinger, & Lawrence, 2005.]

In attempting to replicate the finding of extraordinary performance, one difference in methodology is that this investigation used two-person groups (dyads) instead of the six-person groups of the previous study. One reason for using dyads is that the ideal group model (Sorkin & Dai, 1994) is best specified for dyads. In particular, with dyads, the correlation between members' responses is a Pearson correlation coefficient instead of the intraclass correlation. Because of the nature of the intraclass correlation, it might be an inflated value for r in the ideal group model. However, it can be shown that the intraclass correlation for groups larger than two is equivalent to the Pearson correlation for dyads.

In attempting to replicate the finding of extraordinary performance, all other aspects of the study were similar to the study described directly above. The same video-taped job interview was used. The same recognition items were used for both the initial and second memory test. The instructions were the same, including the emphasis on the groups agreeing on what was the correct answer. Given this replication of the earlier study, many of the predictions were the same. It was predicted that the d' of dyads would be greater than that of comparable individuals. Likewise, dyads should have higher proportion of correct responses. Similar to the previous study, it was also predicted that dyads and individuals would not differ in the β (beta) value indicating bias toward making false alarm responses relative to misses.

One additional prediction for this study is that repeated testing of the recognition memory should influence the β parameter assessing bias. In particular, the second memory test should be more biased toward false alarms ($\beta < 1.0$) because of interference from the initial memory test (Bjork, 2001). When participants respond to the initial memory test, they are exposed to statements that are not true based on the stimuli presented. However, when the participants respond to the same memory test the second time, they will remember seeing the statement previously (Burns & Gold, 1999). Therefore, they may be biased toward responding that the statement was true when it is not. This false alarm of saying a false statement is true will then be higher for the second test. When there are more false alarms than misses, β is lower than when false alarm and miss errors are equivalent ($\beta = 1.0$). Thus, a unique prediction for this study is that participants would have lower β values (farther from 1.0) on the second test than the initial memory test. Because both dyads and individuals will be exposed to the initial memory test, it is not expected that this decrease in β will differ for the two conditions.

The primary purpose for this study was to test the motivation explanation for the extraordinary levels of group performance relative to those predicted by the ideal group model. The main premise was that becoming involved in group task performance motivates the participants such that they exhibit better memory for the material presented. In this way, the memory measures for the initial test do not reflect the actual sensitivity for the material (*d'*). If this was the case, then if participants were motivated to perform better, the *d'* on the memory test should increase. As a consequence, it could be argued that the extraordinary performance observed in the earlier study might result from increased motivation that arises in interacting task-performing groups.

As noted in the discussion of the previous study, some have speculated that engaging in group interaction may motivate group members. One reason that members might become more motivated to perform well on the task is that they feel a greater sense of responsibility when they are a member of a group (Wallach, Kogan, & Bem, 1964). That is, as a function of group membership, group members feel a sense of responsibility to other group members. Moreover, being assigned a task as a group makes members recognize that they will be held accountable (Tetlock, 1985) to other group members during interaction. Group members might want to insure that the reliance other group members' have in them is warranted. Thus, group members may try harder to insure that the group succeeds. Consequently, even telling individuals that they will be interacting with others who will be relying upon them for the collective performance should increase motivation on a task. Asking the individuals to perform as a group should also influence the motivation to perform well in dyads. These arguments suggest that groups provide conditions that serve as a social incentive for group members to do well on tasks.

The potential for a social incentive for group members could lead to the pattern of performance reported in the previous study. In that study, individuals performed the initial memory test on their own without knowledge that they would later perform the task as a group. These eventual group members did not differ from those in the individual condition. However, when the group members performed the task as a group they achieved higher levels of performance than those of individuals on the second memory test. If group membership led the members to have an incentive to remember the material better, then members' *d'* would have been higher. As a consequence of better responding by the group members, the group performance would be much higher. That is, the predictions from the ideal group model, by relying on the *d'* from the initial test, underestimated the members contributions to the group. In this way, the observed group performance appeared to be extraordinary relative to the standard established by the ideal group model.

The current study tested the potential influence of a social incentive. One condition involved telling the eventual dyad members that they would be performing the memory task with their partner and that their partner would be relying upon them to help attain high levels of performance. If this social incentive influences task performance, then participants in this condition should have higher *d'* values relative to control conditions. Moreover, the dyad members subsequently performed the memory task as interacting dyads. This allowed a determination of whether interaction of the members had an independent or critical influence on the social incentive. Interacting with a partner on the memory task might also increase performance, so these social incentive dyads were compared to the performance of a control dyad condition. Forming a dyad and actually interacting on the task might exaggerate the influence of anticipated interaction. Both the anticipation and actual interaction among members of a dyad were included in the current study to determine if social incentives might contribute to enhanced memory performance of groups relative to comparable individuals.

Motivational influences separate from social ones might also enhance group performance. Any motivational influence might have an impact similar to social incentives on individual *d'* and subsequently group performance. Consequently, this current study also included a monetary incentive condition that directly manipulated motivation to do well on the memory tests. Monetary incentives are known to be among the most influential ways of increasing task performance (Locke, Feren, McCaleb, Shaw, & Denny, 1980). Therefore, monetary incentives might be able to motivate individuals to do well on a memory task. Individuals who are motivated to do well on the task might provide more accurate estimates for the ideal group model resulting in more appropriate predictions for the group performance. So, the current study includes a monetary incentive condition in which the eventual dyad members were told they would receive a sizeable bonus if they were among the best performers on the initial memory test. Moreover, these dyad members were offered another bonus if their dyad was among the best performers on the second memory test. These monetary incentives can help determine if monetary incentives can serve as motivators for higher levels of performance. This study can also help determine whether group interaction moderates the influence of monetary incentives on memory performance. The performance from monetary incentive condition will be compared to that of a control condition that receives no incentives as well as the social incentive condition. In conjunction, these different incentive conditions provide a means to test a variety of motivational influences that might enhance the performance of individuals and dyads on recognition memory tests. The details for this study follow.

Method

Participants and Design

Participants were 188 undergraduate students at North Dakota State University who received extra-credit in their lower-level psychology classes. A slight majority were female. These students were randomly assigned to one of four conditions. The individual comparison ($n = 28$) and dyad control ($n = 52$) conditions did not receive a motivation manipulation whereas the dyad with social incentive ($n = 56$) and dyad with monetary incentive ($n = 52$) conditions received specific motivational instructions. All participants completed the initial memory test individually without interaction. Participants in the dyad control, dyad with social incentive, and dyad with monetary incentive conditions completed the second memory test with their partner to produce one set of responses. The individual comparison condition participants performed a filler task and then completed the second memory test again individually.

Memory Materials and Tests

Participants watched the same mock job interview used in prior collaborative memory research (Hinsz, 1990). The interviewee answered general questions about his college and work experiences as well as his aptitude for management. The memory tests involved 60 true-false recognition memory items that could be unambiguous answered correctly from close attention to the interaction presented in the video-recording. Half of the questions were answer correctly as true and half as false. The initial and second memory tests were identical except for five unique items distributed throughout the test. The second memory test had five unique items which were substituted for five items that were then unique on the initial test. Consequently, there were 55 identical questions on the initial and second tests in the same locations in the item sequence.

Procedure

Participants arrived at a waiting room prior to the beginning of the experimental session where they were met by an experimenter who escorted them to one of three experimental rooms. If two participants were randomly assigned to a room they became one of the dyads, otherwise an individual was escorted to a room and completed the experiment alone as an individual in the comparison condition.

After being welcomed to the experiment, the participants were told that they would be watching a videotaped mock job interview of a former student. They were told to pay attention to the video because they would be asked later to complete a memory test about the material presented in the video. Participants in the social incentive condition were also told that "you will eventually also be working with the other person in your room to answer these questions and your partner will be relying on you to help give the most accurate answers to these questions." The students in the monetary incentive condition were told that "you have a chance to make some money if you give among the most accurate answers to the true-false recognition items. That is, the students who are among the top 10% of students who answer these questions most accurately will receive $20 dollars." The individual comparison and dyad control conditions did not receive any incentive instructions.

When the video finished, the participants were asked to retrieve a booklet of the memory test items and computer-readable form upon which to mark their responses. The participants were asked to read each of the true-false questions carefully and mark whether they thought the item was true or false from what they watched on the video. They were then to respond indicating how confident they were in their true-false response. The participants in the social incentive condition were reminded "that later you will be asked to complete the true-false recognition items and the other person in your room will be relying on you to help give the most accurate answers to these questions." Participants in the monetary incentive condition were told to "Remember that you have a chance to make $20 if you have among the most accurate answers to the true-false recognition items." The participants were then told to answer all 60 items individually and without discussion.

When all participants in the room had finished the initial memory test, they were asked to place those materials out of the way. Then the participants in the dyad conditions were asked to once again complete the (second) memory test. Each dyad was asked to retrieve a pair of booklets listing the memory items and one computer-readable form that they would use to give their dyad's responses to the memory items. The dyads were told that we wanted them "to jointly respond as a dyad. Now, you may reach your joint responses and judgments in any way that you wish, but your dyad's final responses must represent your dyad's collective opinion about if the question is true or false and the degree of confidence you jointly have in your dyad's true-false response. During your discussion of these various questions, make sure that each person has the opportunity to speak, and make sure that you give due consideration to the opinions of the other member of your dyad." Participants in the monetary incentive condition were also told that "your dyad has a chance to make $40 dollars if your dyad is among the top 10% of dyads with the most accurate answers to the true-false recognition items. That is, if your dyad is among the most accurate dyads, you will receive $40 dollars, $20 per member of your dyad for being the most accurate." If the dyad members did not have any questions, they were allowed to complete their memory test at their own speed.

After completing the initial memory test, participants in the individual comparison performed a filler task to account for the longer time it would take for the dyads to complete the second memory task. The filler task involved responding to nine problems regarding the use of base-rate and test-reliability information in medical decisions (Hinsz, Heimerdinger, Henkel, & Spieker, 2005). These individuals then completed the second memory test again alone and without incentive instructions.

When all participants completed the second memory test, they were asked to place the materials associated with the test aside, and then retrieved a questionnaire. This questionnaire included a set of items that assessed how attentive, motivated, and committed to do well the students were on the second memory test. Upon completion of this questionnaire, the students

were debriefed and thanked for their participation. Students in the monetary incentive condition provided contact information to be used if they were among the top performers. The top performers were later determined and they were given the money promised.

Results

Recognition Memory Accuracy

Table 1 contains the proportion of items answered correctly by individuals and all dyad members (collapsing across motivation conditions) on the initial recognition test and second recognition test. Table 2 contains the proportion of items answered correctly by each dyad condition on the second recognition test. These analyses involving the proportion correct data were conducted following an arcsine transformation. The individual and dyad conditions did not differ on the initial recognition test, $F(1,186) = .00$, *ns*. The dyads and individuals did differ in the accuracy of their responses on the second recognition test, $F(1,106) = 11.90, p < .001$, partial $\eta^2 = .10$. Dyads gave a higher proportion of correct responses than did individuals on the second recognition test.

Table 1

Mean Values and Standard Deviations of the Memory Measures for Recognition Test 1 and 2

	Initial Recognition Test		Second Recognition Test	
Dependent Variable	Individual *M* (SD)	Dyad Member *M* (SD)	Individual *M* (SD)	Dyad *M* (SD)
Proportion Correct	0.673 (0.078)	0.674 (0.076)	0.674 (0.079)	0.725 (0.063)
d′	0.973 (0.485)	0.984 (0.464)	0.983 (0.508)	1.276 (0.405)
β (Beta)	0.843 (0.153)	0.874 (0.225)	0.819 (0.272)	0.793 (0.247)

Table 2

Mean Values and Standard Deviations of the Memory Measures for Recognition Test 2 by Dyad Condition

Dependent Variable	Control Dyads *M* (SD)	Social Incentives Dyads *M* (SD)	Monetary Incentives Dyads *M* (SD)
Proportion Correct	0.685 (0.071)	0.689 (0.081)	0.699 (0.074)
d′	1.037 (0.427)	1.081 (0.510)	1.126 (0.454)
β (Beta)	0.853 (0.234)	0.841 (0.257)	0.848 (0.215)

A repeated measures analysis of variance (ANOVA) of dyad condition by test revealed a significant effect of accuracy of responses on the initial versus second test, $F(1, 72) = 27.64, p < .001$, partial $\eta^2 = .28$. The proportion correct on the initial recognition test ($M = .67$) was significantly less than the proportion correct on the second recognition test ($M = .73$). However, there was no effect of incentive condition on memory performance which is inconsistent with expectations, $F(2, 77) = 1.55, p > .21$. The different incentive conditions did not influence the performance of dyads or dyad members. Rather, the primary effect was that dyads had better memory performance than individuals, which is consistent with research on collective memory performance (Clark & Stephenson, 1989; Hinsz, 1990; Hinsz et al., 1997).

Signal Detection Analyses

Based on the proportion of hits and false alarms, the d' and β measures of signal detection were calculated for the individual and dyad conditions on the initial and second memory test. Summary statistics for these values for individuals and all dyads on both recognition tests are given in Table 1. Table 2 contains the summary statistics for the motivation conditions on the second recognition test.

Memory Strength, d′. One-way ANOVAs were conducted on d' values for each individual and dyadic condition on the initial test and for individual and dyad conditions on the second test. Results of these analyses revealed that the individual control condition and the dyad conditions did not differ in the initial recognition test, $F(1,186) = 0.01$, *ns*. However, significant differences in memory strength were found in the second recognition test performance, $F(1,106) = 9.44, p < .01$, partial $\eta^2 = .08$. Dyads were significantly more sensitive to the correct answer in their responses to the second recognition test than were individuals. A 3 (dyad control/social incentive/monetary incentive) by 2 (recognition test) repeated measures ANOVA was conducted on responses to the second recognition test to examine differences in d'. No differences were observed in second recognition test performance among the motivation conditions, $F(2,160) = 0.77$, *ns*. Analyses also revealed that d' on the initial recognition test ($M = .98$) was less than it was on the second recognition test ($M = 1.28$), $F(1, 72) = 22.79, p < .001$, partial $\eta^2 = .24$.

Decision Strategy, β. One-way ANOVA tests were conducted on β values for the individual and dyad conditions on the initial test and for individual and dyadic conditions on the second test. As predicted, results of these analyses revealed that the individual condition and the dyad conditions did not differ in the initial recognition test, $F(1,186) = 0.48$, *ns*., nor did they differ in the second recognition test, $F(1,106) = .21$, *ns*. A 3 (dyad control/social incentive/monetary incentive) by 2 (recognition test) repeated measures ANOVA was conducted to evaluate the hypothesis concerning differences in β over the tests. A significant difference in β over recognition tests was observed, $F(1, 72) = 6.39, p < .01$, partial $\eta^2 = .08$, with β being higher on the initial recognition test ($M = .87$) than on the second ($M = .79$).

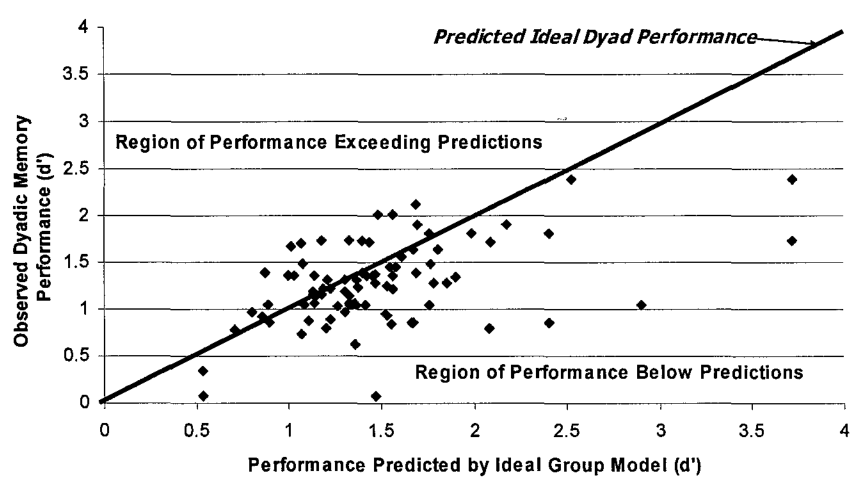
Ideal Group Model Analysis

To calculate the ideal group d' for this sample of dyads across all conditions, measures of the parameters were needed. These parameters and the resulting analyses are contained within Table 3. Consistent with the result reported directly above, there were insubstantial differences in the parameters for the different incentive conditions, so analyses focused on the dyads conditions collapsing across all motivation conditions. Results suggest that dyads responding to the recognition memory tests were on average integrating the information they had available at suboptimal levels. The dyads achieved 87% of efficiency. These dyads did suffer from coordination, process, motivation, or other losses (Steiner, 1972) when integrating the information. Further break down by motivation condition suggests that none of the conditions, including those conditions designed to enhance motivation, were integrating the information they had available at optimal levels as was reported in the previous study.

Table 3							
Ideal Group Parameters and Analyses							
	Group Size	ICC Among Initial Responses	Mean d' of Dyad Members	Variance in Dyad Members d'	Ideal Group predicted d'	Observed Group d'	Efficiency (η)
All Dyad Conditions	2	0.177	0.984	0.216	1.473	1.276	87%
Control	2	0.187	0.938	0.169	1.378	1.234	90%
Social Incentives	2	0.183	1.012	0.297	1.568	1.219	78%
Monetary Incentives	2	0.156	1.000	0.180	1.468	1.378	94%

To examine the performance of the dyads relative to the ideal group model, analyses were also conducted by dyad. The dyad's d' values for their observed performance ranged from .08 to 2.39 whereas the predicted ideal group d' values for these dyads ranged from .53 to 3.72. A scatter plot of the observed versus predicted d' values is presented in Figure 2. The η values for these dyads ranged from 0% to 270%, indicating that although on average the dyads performed at suboptimal levels ($M = 92\%$), 31% of the dyads exceeded the predictions from the ideal group model. These rates of dyad performance exceeding ideal group model predictions were lower than those reported for the six-person groups above, Fisher's exact $p < .03$. Consequently, although portions of these result replicate those reported above, other results differed in important ways.

Figure 2. Observed Dyad Performance Relative to Predicted Ideal Group Performance



Discussion

The results of this study provide important results regarding the replication of extraordinary performance in interacting groups and the potential for motivation to influence performance on the recognition memory task. This study replicated previous research that shows collective performance to be superior to that of individuals on a memory task (Clark &

Stephenson, 1989; Hinsz, 1990; Hinsz et al., 1997). However, although superior to that of individuals, the dyad's memory performance did not meet or exceed that predicted by the ideal group model. In contrast to predictions, the results of this study demonstrated no motivational influences on recognition memory performance. Consequently, this study adds significantly to research attempting to replicate and explain the extraordinary performance reported above.

The results of this study are quite consistent with much of the research examining memory performance in groups. First, collective performance exceeded that of individuals who were treated similarly. Second, the collective performance, although superior to that of individuals, was still suboptimal relative to standards that can be constructed for models of optimal performance. Third, the tendency of the recognition memory test to produce proportionally more false alarms than misses was replicated. Fourth, the findings from signal detection analyses were consistent with other measures of memory processes. In conjunction, these results suggest that the existing findings for collective memory performance are quite robust across group sizes, materials to be remembered, and measures used for memory performance and processes (Hinsz et al., 1997).

This study did not replicate the finding of extraordinary performance relative to the ideal group model as reported above and demonstrated in Figure 1. The failure to replicate could arise for a number of reasons. Perhaps dyads are different from the six-person groups that were reported in the study above. The previous study and this one differ in how the dependence in member responses was incorporated in the predictions of the ideal group model. That is, this study used the Pearson correlation coefficient while the previous study used the more general intraclass correlation. Alternatively, perhaps the interaction processes of dyads are substantially different from those of six-person groups. The complexity of interactions with more co-actors may be necessary to produce the extraordinary performance that was observed above. Or, perhaps some of the values used to make the ideal group model predictions are influence by group size. That is, collective performance meets or exceeds performance predicted by values derived from six-person groups while the values used for dyads result in judgments of suboptimal performance. Nevertheless, it is clear that a much better understanding of group performance and the way the ideal group model predicts it is needed. The study that follows, which involves a simulated AWACS task, will hopefully improve the understanding of team performance and model predictions of that performance.

The results of this study provide evidence that motivational influences do not play a role in extraordinary performance reported above. Although there are some important differences between the previous study and this one, primarily group size, the similarities are widespread. Consequently, the tests of the influence of the social and monetary incentives on memory performance can be considered directly. Although there was some speculation that group interaction might enhance member motivation, this study found no support for that speculation. The social incentive manipulation did not increase members' memory performance on the initial test. The speculation that interaction might influence the performance of dyads relative the control dyad condition was not supported. Similar results were found for the monetary incentive condition. The chance to earn money for being among the most correct respondents did not lead the dyad members or the dyads to achieve better memory performance. The lack of differences between the control condition and the social and monetary incentive conditions suggests that participants are sufficiently motivated to do well on the memory tests such that no more incentive is necessary or effective. Alternatively, given the strong cognitive basis of memory performance on this task, perhaps motivational influences play a non-significant role once some basic level of motivation is applied to the task. Nevertheless, this study found no influence of the

incentive motivational manipulations on memory performance, implying that motivation is unlikely to be responsible for the extraordinary group performance reported in the study above.

This study did reveal support for the hypothesis for more false alarms in the responses to the second memory test. As predicted, the bias measure (β) was significantly lower on the second test than on the initial test. This difference indicates that the exposure of the participants to the incorrect statements on the initial test may have given the participants a false sense of familiarity with the material. This memory confusion could lead the respondents to indicate that a statement that wasn't accurate from the video recording would be perceived as correct on the second test because they had seen the statement on the initial test. This would result in the higher numbers of false alarms observed for the second test in this study and the corresponding β value. This result enhances the understanding of the pattern of errors that occur in both the individual and collective memory performance of the studies.

In summary, this study adds substantially to the understanding of group memory performance and the potential for extraordinary performance that was discussed above. First, the extraordinary performance described in the previous study is not robust enough to be replicated with two-person groups using the same materials and methods. This might occur because the pattern of group performance is moderated by group size. Or, the different means of constructing the parameters used in the ideal group model are greatly influenced by group size. Second, if extraordinary group performance relative to the predictions of the ideal group model does arise, it is unlikely to result from increased motivation that arises in groups. More likely, as described in group performance literature (Sheppard, 1993), groups contribute to reductions in motivation among group members. Third, the importance of replicating research is again demonstrated. This study raises important questions about the findings of the previous study. Moreover, by adding some interesting motivational manipulations and further examining measures (β), the understanding of group performance on recognition memory items was enhanced. The following discussion considers another conceptual approach to group recognition memory performance which can further contribute to an understanding of group performance.

9. Social Decision Schemes and Team Decision Making

Combinations-of-contributions theory provides a general perspective for how teams would respond to recognition memory items. A team's response would reflect the memories the team members bring to the discussion (i.e., contributions) and some combination of these member memories. To understand and provide an adequate model of weapons director teams' memory processes, it would be useful to assess the members' memories for the material, and then test which combination of member memories best describes the way the teams arrived at their collective responses. Social decision scheme theory (Davis, 1973, 1982) provides a direct means of addressing these questions. Fundamentally, the social decision scheme (SDS) model aims to uncover the ways in which the responses of team members can be combined into the team's response. This involves three primary considerations: the distribution of the team members' responses, the aggregation principle that combines these responses (termed a decision scheme), and the means of testing the adequacy of the decision schemes in predicting a sample of observed team responses (model testing).

The general SDS model assumes that each team member, and subsequently each team, selects one of a set of possible responses (i.e., true or false). A single team of m individual members can array themselves over a set of possible *distinguishable distributions* of responses. For example, in the previous study with six members in the group, there would be seven distinguishable distributions of how the members arrange themselves in terms of having correct

(C) and incorrect (I) responses to the recognition memory items. Table 4 shows the potential distributions of member preferences in the left-hand column: 6-0, 5-1, 4-2, 3-3, 2-4, 1-5, and 0-6. In addition, the next two columns to the right in Table 4 show the proportion of groups having that distinguishable distribution that answered the questions correctly or incorrectly. Note that for these groups, as the number of group members initially having the correct answer increases, so does the probability that the group will respond correctly. However, also note that groups did not respond correctly just because one or more group members initially had the correct answer. Consistent with the research on conformity (Asch, 1956), if one lone correct member faced a unanimous majority of members (five) that were incorrect, the majority persevered in the group judgment more often that the correct individual. It could be argued that the groups in these situations were not responding in an optimal fashion. But, it is also important to recognize that even when none of the group members initially responded correctly, the group was able to get the correct response one third of the time. This reflects extraordinary performance that is relevant to the tendency for group performance to exceed expectations.

Table 4. Observed Distribution of Group Recognition Responses Across All Items (Hinsz, 1990).

Member Responses (C, I)	Group Correct	Response Incorrect
6,0	.98	.02
5,1	.94	.06
4,2	.88	.12
3,3	.77	.23
2,4	.61	.39
1,5	.41	.59
0,6	.33	.67

To provide a conceptual understanding of how teams might combine their member responses, a *social decision scheme* is proposed (Davis, 1973; Stasser, 1999). A social decision scheme is a rule or procedure that conceptually combines (usually in algebraic fashion) the various individual responses into a single team response. Decision schemes can be constructed to represent a variety of different social processes hypothesized to underlie team judgment. Research following the SDS approach aims to identify plausible decision schemes that might describe the implicit decision rule that may exist for a sample of teams. A host of decision schemes can be considered for team responses to memory items (Hartwick, Sheppard, & Davis, 1982; Hinsz, 1990). Hinsz (1990) proposed six specific decision schemes and found that the plurality correct, incorrect otherwise decision scheme best fit the data gathered (see Table 4). This decision scheme suggests that if three or more group members initially responded with the correct response, then the group was likely to respond correctly. However, if a minority (two or fewer) of members responded correctly, then the group was likely to respond incorrectly. Another decision scheme that suggested an optimal process was found not to be a plausible account for the data. These analyses help highlight the model testing aspects of social decision scheme analyses.

The pattern of member and group responses depicted in Table 4 can inform us much about potential processes by which group responses are reached. It also complements the understanding of group performance provided by the ideal group model. It is clear that not all groups perform at optimal levels because if they did, then anytime one member was initially correct, the group should respond correctly. However, it also points out that groups can perform better than expected. In particular in the 0-6 condition, when all members were initially incorrect but the group responded correctly, that any groups responded correctly exceeds general expectations of how groups would operate. Consequently, an inspection of Table 4 suggests that under differing conditions, groups might reach suboptimal performance while other groups achieve supraoptimal performance. Clearly, further analyses of group performance from the ideal group model and social decision scheme theory perspectives could help us understand why teams

such as weapons directors might achieve optimal levels of performance and perhaps exceed our general expectations of their teams' performance.

Social decision schemes analysis of the second study reported above might also illuminate important processes of team performance on memory tasks. Table 5 presents the general pattern of dyad member and dyad responses. Because there were only minor differences between the control, social incentive, and monetary incentive conditions, Table 5 reflects responses collapsed across the incentive conditions. Note that similar patterns of results are found for the dyads as were found for the six-person groups reported above. In particular, as the number of initially correct members increases, the proportion of dyads answering correct increases as well. When both members of the dyad initially respond correctly, the vast majority of dyads answered correctly, although there were still occasions when the dyads responded incorrectly. Also, when both members initially responded incorrectly, the dyads were likely to respond incorrectly, but a sizable set of responses moved to the correct response. In addition, similar to the 3-3 condition in the six-person groups, groups in the 1-1 dyad condition were twice as likely to respond correctly as incorrectly. So, similar to findings from the earlier study, group performance on this memory task was influenced by the number of members initially favoring an alternative, but whether the members were correct or incorrect also had an impact. These general patterns should be highly informative for AWACS weapons director teams that also face memory tasks in their team performance. We now turn to a further consideration of AWACS weapons directors' team performance and the implications of these social decision scheme and ideal group model results for these teams.

Table 5. Observed Distribution of Dyad Recognition Responses Across All Items and Conditions.

Member Responses (C, I)	Dyad Response Correct	Dyad Response Incorrect
2,0	.92	.08
1,1	.67	.33
0,2	.17	.83

10. Team Performance among AWACS Weapons Directors

The studies described so far have focused on memory performance in interacting groups. This research is relevant to weapons director teams because memory is a critical cognitive process in the multiple tasks performed by these teams. Whenever weapons directors lose access to the screen display information, they must rely upon their memory to perform their tasks. One way weapons directors can effectively respond to this situation is to use collaborative strategies to enhance the memory performance of their team. This report aims to describe experiments that explore the impact of collaboration on memory performance. This next experiment to be discussed considers collaboration of simulated weapons directors in a memory task on a synthetic task environment representative of AWACS weapons directors (Fahey et al., 2000).

The memory performance of simulated weapons director teams can be conceptualized from the ideal group model and social decision scheme perspectives described above. Moreover, from these previous studies, predictions about potential levels of performance can be derived from both the ideal group and social decision scheme theoretical models. These theoretical models provide a means for generating benchmarks for differing levels of performance. Benchmarks for optimal performance describe what would be required for the teams to achieve maximum efficiency and optimum use of member contributions. In order for research to improve the performance of weapons director teams, it will be beneficial to know the degree to which interventions lead teams to approach optimal levels of performance. Both the signal detection theory based ideal group model and social decision scheme theory provide conceptual bases upon which to specify optimal levels of performance given weapons directors' individual

performance levels.

The theoretical approaches applied in this experiment also provide a basis for considering non-optimal levels of performance. It might be important to recognize how poorly a team performs relative to sub-optimal benchmarks for performance. The models considered can define lower bounds of performance that would be expected if specific team processes occurred in weapons director teams. Consequently, it might be possible to identify processes of team memory performance that are leading the weapons directors to fail in their task performance. In addition, benchmarks for lower-bound levels of performance can be constructed which would indicate how poorly teams might perform on a memory task if the teams used wholly inappropriate processes. The models can also be used to describe the processes of collaborative memory performance in weapons director teams and establish descriptive benchmarks for normal collaborative memory performance of weapons directors. This experiment explores the degree to which the performance of weapons director teams approaches the various benchmarks for performance.

Although this experiment focuses on an understanding of optimal and suboptimal performance of weapons director teams on memory tasks, it also provides an opportunity to replicate the levels of performance that exceed optimal as predicted by the models (i.e., supraoptimal). This experiment considers how supraoptimal levels of performance might arise for weapons director teams and provides a way of assessing supraoptimal performance. If, under the conditions of this experiment, supraoptimal levels of performance were reached, it would transform the ways in which team performance is considered. Team researchers would have to totally revamp their thinking about the processes and mechanisms involved in maintaining and attaining effective team performance.

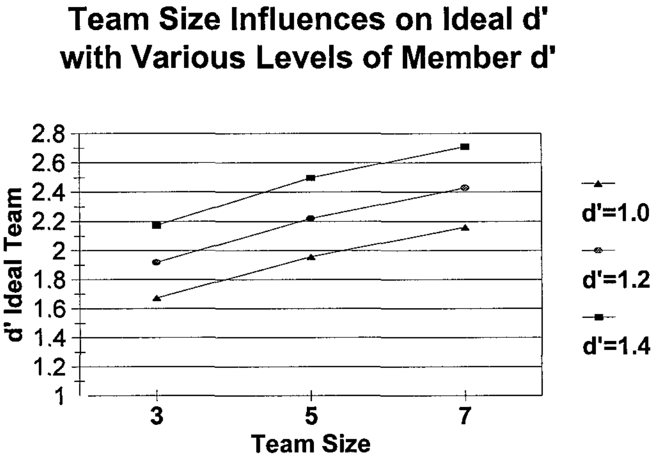
An additional aspect of this experiment is that it involves teams of different sizes. One issue that the Air Force faces is reductions in force that may necessitate changing force structure. One consequence of the reduction in force is that various Air Force operational teams might have to perform the same mission with fewer members. Current Air Force AWACS weapons director teams usually consist of five members, typically a senior director and four weapons directors. This experiment examines the potential implications of reductions in team size, particularly as it influences collaborative performance on a recognition memory test for material presented on the weapons director's displays. The ideal group and social decision scheme models make specific predictions regarding the impact of reductions of team size on performance. The results of the studies described above involving six- and two-person groups suggest that complex changes in team processes and outcomes arise when teams of different sizes are assigned the same task. Consequently, it is important to determine if reductions in weapons director teams from five members to fewer (e.g., three members) will influence the processes and performance of those teams. In this experiment, five-person weapons director teams will be compared to three-person teams to determine the influence of team size on levels of performance and performance relative to theoretically-derived benchmarks.

This experiment has a conceptual foundation based upon the ideal group model and social decision scheme theory. The implications of these theoretical approaches will each be considered next as they relate to the conditions produced in the experiment.

Ideal Group Model

Recall that Sorkin and colleagues proposed an ideal group model based upon signal detection theory properties. The ideal (optimal) d' for a team is calculated based on four parameters: team size, the mean detectability of the team members, the variance of the member detectability indices, and the intercorrelation of the responses of the team members (ICC). Based on these four parameters, the ideal team d' is calculated according to Equation 1 (see above).

From this equation, it can be discerned that ideal team d' increases as team size increases. Figure 3 illustrates the predicted differences for three, five, and seven person weapons director teams with changes in d', assuming var(d')=.20, and ICC=.25 (based on the data presented above from Hinsz, 1990). Simulations indicate that changes in variance in member d' and the ICC will generally produce parallel lines in the figure for teams of 3, 5, and 7 members.



This ideal team d' model also suggests the patterns of responses that might be expected if weapons directors are asked to respond to recognition items as members of teams of different sizes. Similar to ideal team d', observed team d' should increase as team size increases, the observed mean d' and variance of member d' increase, and as ICC decreases. The members of weapons director teams can undergo changes that will lead to changes in mean d' (e.g., increased training and heightened motivation), variance in d' (e.g., differences among members' experiences on the task) and the intra-class correlation (e.g., dissimilarity of members' background and the way they received information). Although these are interesting issues for consideration of the performance of weapons director teams, this experiment focuses on differences in team size (i.e., 3 and 5 member teams). Among the sample of teams that will be tested, it is expected that there will be variation in mean d', the variance of member d', and ICC, thus allowing an examination of the ability of the ideal group model to account for team performance on the recognition memory task.

In addition to the identification of the ideal team d', this model also provides a way to examine performance relative to the ideal team performance (i.e., optimal). Sorkin argues that the degree a team d' approaches its ideal (optimum) reflects the efficiency of the team in using its available information. The efficiency of the observed team is specified as η. A team that reached the level of performance specified by the ideal d' would have an efficiency of 100%. This analysis provides a definition of optimal performance by providing a concise index of the efficiency of weapons director teams' performance. As demonstrated in the studies above, this efficiency index also suggests a way to demonstrate supraoptimal performance if it arises. Consequently, this experiment examines the efficiency indices of the weapons director teams as well as the ideal d' for the teams.

Social Decision Schemes Theory

The social decision scheme approach described above can also be applied to the issues considered in this experiment. The same three issues of distinguishable distributions, plausible decision schemes, and model testing of the predictions from decisions schemes against observed responses apply in situations of simulated weapons director teams responding to recognition memory items. It can be shown that there are four distinguishable distributions for a three-person team responding to true-false questions, and six distinguishable distributions for a five-person team. The possible distributions of correct and incorrect responses for a five-person team would be 5-0, 4-1, 3-2, 2-3, 1-4, 0-5 (see Table 6). The distributions for a three-person team would be 3-0, 2-1, 1-2, 0-3 (see Table 7). Tables 6 and 7 present the initial set of decision schemes that will be considered in this experiment. Each decision scheme reflects a set of processes that teams might follow when trying to choose a response to memory items.

Table 6. Hypothesized Decision Schemes for Memory Responses of Five-Person Weapons Director Teams.

Single Correct Decision Scheme

Member Responses	Team Correct	Response Incorrect
5,0	1.00	.00
4,1	1.00	.00
3,2	1.00	.00
2,3	1.00	.00
1,4	1.00	.00
0,5	.00	1.00

Pair Correct Decision Scheme

Member Responses	Team Correct	Response Incorrect
5,0	1.00	.00
4,1	1.00	.00
3,2	1.00	.00
2,3	1.00	.00
1,4	.00	1.00
0,5	.00	1.00

Majority Correct Decision

Member Responses	Team Correct	Response Incorrect
5,0	1.00	.00
4,1	1.00	.00
3,2	1.00	.00
2,3	.00	1.00
1,4	.00	1.00
0,5	.00	1.00

Strong Majority Decision Scheme

Member Responses	Team Correct	Response Incorrect
5,0	1.00	.00
4,1	1.00	.00
3,2	.00	1.00
2,3	.00	1.00
1,4	.00	1.00
0,5	.00	1.00

Worst Case Decision Scheme

Member Responses	Team Correct	Response Incorrect
5,0	.50	.50
4,1	.00	1.00
3,2	.00	1.00
2,3	.00	1.00
1,4	.00	1.00
0,5	.00	1.00

Supraoptimal Decision Scheme

Member Responses	Team Correct	Response Incorrect
5,0	1.00	.00
4,1	1.00	.00
3,2	1.00	.00
2,3	1.00	.00
1,4	1.00	.00
0,5	.50	.50

Table 7. Hypothesized Decision Schemes for Memory Responses of Three-Person Weapons Director Teams.

Single Correct Decision Scheme

Member Responses	Team Response Correct	Team Response Incorrect
3,0	1.00	.00
2,1	1.00	.00
1,2	1.00	.00
0,3	.00	1.00

Majority Correct Decision Scheme

Member Responses	Team Response Correct	Team Response Incorrect
3,0	1.00	.00
2,1	1.00	.00
1,2	.00	1.00
0,3	.00	1.00

Supraoptimal Decision Scheme

Member Responses	Team Response Correct	Team Response Incorrect
3,0	1.00	.00
2,1	1.00	.00
1,2	1.00	.00
0,3	.50	.50

Worst Case Decision Scheme

Member Responses	Team Response Correct	Team Response Incorrect
3,0	.50	.50
2,1	.00	1.00
1,2	.00	1.00
0,3	.00	1.00

The first matrices in Tables 6 and 7 illustrate the optimal decision scheme termed single correct. In this decision scheme, if any one of the team members knows the correct response before the discussion, the team is predicted to respond with the correct response. This decision scheme incorporates the notion of the optimal team process as described by Lorge and Solomon (1955) and Laughlin (1980). Although a single-correct decision scheme would produce optimal performance for teams responding to memory items, research has rarely found teams to achieve this level of performance. Generally, a second advocate favoring the correct response is required for a team to respond correctly (Davis, 1982; Hinsz, 1990; Laughlin, 1980). The pair-correct decision scheme reflects this notion that the team is most likely to chose the correct response if two members know the correct response. In this case, a member with an initially correct response would have some support from within the team to demonstrate the correct response. As conformity research indicates (Asch, 1956), a person is more likely to standup to a strong majority and advocate the correct response if an additional team member also supports the correct response.

Research on team decision making indicates that a pair of team members may not be sufficient to dominate in reaching a response. Rather, in the cases in which there is no correct response, the decision process appears to follow a process similar to majority rules (Davis, 1982; Laughlin, 1996). It seems plausible that teams might also reach a correct response if they have a majority of members that initially favored the correct response. This majority-correct decision scheme is provided in Tables 6 and 7 as well. Other decision making research has found that a simple majority does not adequately summarize the processes when the decisions are important or have serious consequences (Davis, 1980, 1992). A strong majority would arise in these cases in which a simple majority provides insufficient social justification or support to take positions that address critical issues or have grave consequences (e.g., applying the death penalty). For some memory items considered by weapons director teams, a strong majority might be required (e.g., there were no hostile aircraft in this sector). Although unlikely for most memory items, it is

important to consider this strong-majority decision scheme to be able to eliminate it as a plausible description of the processes by which weapons director teams respond to memory items.

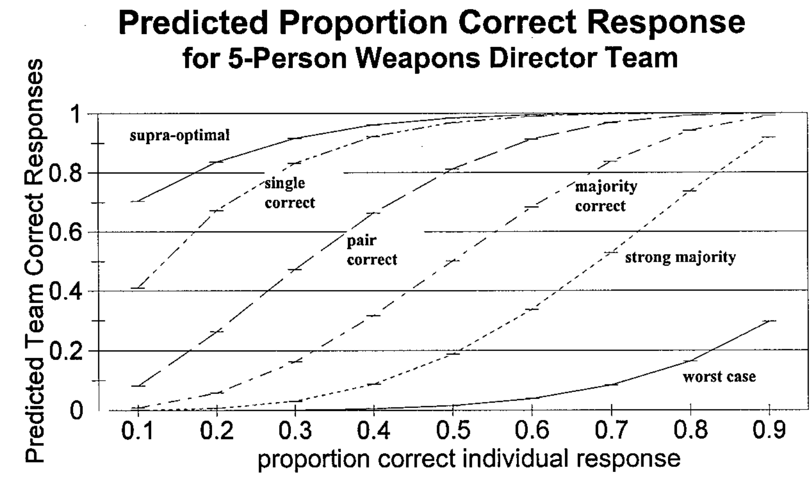
It is also possible to consider a team responding incorrectly even if all the members initially favored the correct response. Although an unlikely decision process, it would reflect the worst-case decision scheme which corresponds with a pessimistic view of team processes (e.g., a camel is a horse designed by a committee; a team must have thought up this idea because no individual alone would be that stupid). If it is presumed that some proportion of teams (e.g., half) will make the incorrect response even if all the members of the team initially favored the correct response, a worst-case decision scheme could be defined (see Table 6). This worst-case decision scheme provides a lower bound benchmark for performance expectations.

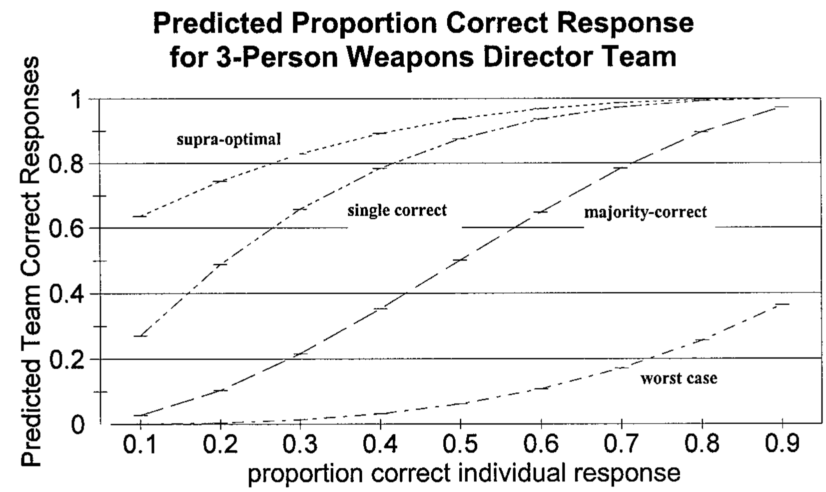
An additional decision scheme will be considered reflecting supraoptimal performance. In this case, the team responds correctly to a memory item even if none of the team members initially responded correctly. This supraoptimal decision scheme suggests that there are specific synergistic benefits of bringing people together as a team to perform a task. These teams are then capable of achieving outcomes that no members of the team might have been able to attain alone (e.g., assembly effect bonus, Collins & Guetzkow, 1964). The supraoptimal decision scheme predicts that teams would be likely to respond correctly some portion of the time (e.g., 50%) even if none of the team members responded correctly to the memory item.

The decision schemes that are applied for five-person teams are more diverse than for three-person teams. For the three-person teams, the decision schemes include the single correct (optimal), worst case, supraoptimal, and majority correct. For the three-person teams, majority correct and pair correct are operationally the same decision scheme. This experiment should be able to differentiate the decision process that best represents these memory responses because the five-person team differentiates between the majority-correct and pair-correct decision schemes. It is also important to note that these decision schemes are described in terms of the more critical correct response. Decision schemes could be constructed with the incorrect responses being important (e.g., majority incorrect). An inspection of the decision scheme matrices in Table 6 reveals that the decision schemes could be labeled in terms of incorrect responses. For example, a majority-correct decision scheme would be equivalent to a majority-incorrect decision scheme for the five-person teams. Consequently, the decision schemes selected and presented in Table 6 provide a broad basis for considering the optimal and actual performance of weapons director teams on a memory task.

The decision schemes presented in Tables 6 and 7 can be shown to make differential predictions regarding the nature of performance on memory items. Figure 4 presents the predicted proportion of correct team responses for the different decision schemes across a range of the proportion of correct initial individual responses. These values are calculated by applying the social decision scheme mathematical formulation for the distribution of responses and the potential decision schemes as described in Tables 6 and 7. Figure 4 demonstrates how the optimal, worst-case, and supra-optimal levels of performance set benchmarks for comparison against the proportion of correct responses that will be gathered in this experiment. The research reported above on memory performance in six-person groups suggests that the decision scheme that will best represent the weapons director teams is between the majority-correct and pair-correct decision schemes. To be able to determine which decision scheme best represents the weapons director teams, the model testing aspect of social decision scheme theory can be applied.

The predicted social decision schemes can be tested by comparing the observed distribution of team responses to the distribution predicted by the proposed social decision scheme using a model testing procedure (Davis, 1973; Kerr, Stasser, & Davis, 1979). A goodness-of-fit statistic can then indicate whether the social decision scheme is an adequate description of the process the team used to make its response. If the predicted distribution of team responses does not differ from the distribution of team responses observed in the sample, the social decision scheme may be considered a plausible description of the process by which the team responses were reached. However, if the two distributions do differ, then that decision scheme can be eliminated as a plausible description of the process for the teams and conditions involved in the study.





The structure of the decision schemes identified in Tables 6 and 7 reflect the some of the processes thought to contribute to the superiority of team performance on cognitive tasks. Information pooling is reflected because any one of the team members may provide the correct response. Thus, teams are more likely to have a member that advocates the correct response. Moreover, error correction arises because a correct response might occur for a team even if a majority of the members initially responded incorrectly (i.e., single or pair correct). In these cases, members that know the correct response are able to lead the team to a correct response, and correct the errors of some of the members. Because of information pooling and error correction, teams are able to improve upon the responses of similarly treated individuals to produce sets of responses that are superior to those of individuals.

The social decision scheme and ideal group model approaches to team performance on a recognition memory task provided bases for establishing benchmarks for performance. In this way, the experiment provides ways of evaluating the influence of collaboration on memory performance. Because this experiment includes an individual comparison condition, it is also possible to examine how team performance relates to individual performance. Given past research, it is expected that both the three and five-person teams should have superior performance relative to individuals. An addition objective of this experiment is to describe the

basic features of the memory performance of weapons director teams. This includes the average level of memory performance (e.g., proportion correct) and other indices used to reflect aspects of the teams' memory processes (e.g., d'). This experiment will also indicate the frequency of different types of errors (e.g., errors of omission and commission) in the memory responses as well as bias in types of errors (e.g., $\exists$). Thus, a number of features of the memory processes of weapons director teams will be revealed as a function of this experiment.

Method

Participants and Design

Participants were 116 students from North Dakota State University with about equal numbers being male and female. Attempts were made to recruit teams of Aerospace ROTC members and members of the local air guard, but few chose to participate. Participants either received extra credit for their lower level psychology classes or were paid a subject fee for their participation. This experiment involves comparisons of weapons directors performing alone with teams of three or five members. There were 26 participants in the individual condition, 15 three-person teams (45 participants), and 9 five-person teams (45 participants). All participants performed the recognition test initially alone. Following that, the individuals performed a filler task, and then performed the memory test again after this delay. These individuals serve as a baseline condition for comparison with the team conditions. Members of the teams performed the recognition memory test the second time as an interacting team.

Task Environment

The Aptima DDD-AWACS synthetic task environment that simulates features of weapons director tasks served as the research platform. While performing their tasks, the participants (i.e., weapons directors) were each seated at a workstation (see Apendix A, Image 1). Participants in the five person teams worked in one room in which the workstations were arrayed against one wall (see Appendix A, Image 2). The participants in the three-person teams worked at a set of three workstations in a different room (see Appendix A, Image 3). The individuals responded to the tasks at one of two workstations in an area set off from the team work stations. Each of these areas was arranged so that the experimenter could monitor the activities of the weapons directors as well as introduce the various procedures that arise during an experimental session. The experimenter controlled and monitored the operation of program and the teams from a control station (see Appendix A, Image 4). Each experimental session involved a series of related events termed a scenario.

A scenario was developed to allow testing of the hypotheses and models proposed for this experiment. Because up to five weapons directors could be involved in the teams, a complex scenario was constructed so that memory processes could be adequately tested (e.g., no ceiling effects). The scenario involved little contact with enemy forces and focused on a humanitarian relief effort in Turkey after it suffered a devastating earthquake. A number of villages and cities were located in the region of responsibility and humanitarian relief (e.g., food and medical supplies) needed to be delivered by air transports. This complex scenario would require the efforts of a number of AWACS weapons directors so that it would be reasonable to compare performance of three and five person teams. Image 5 in Appendix B demonstrates the map grid on the screen display for this scenario at the beginning of time period.

In general, the information displayed to the weapons directors fall into four categories: (1) Aircraft that are under the supervision of the weapons directors on the AWACS (e.g., fighters, helicopters). (2) Targets that are unfriendly, enemy, or of unknown disposition. (3) Background information about the lay of the land and other features of the place and space in the AWACS area of responsibility (e.g., national boundaries, villages). Much of this information is

displayed on a grid with aircraft identified with icons and alpha-numerical labels (see Appendix B, Image 6). In addition, colors are used to indicate the aircraft under a specific weapon director's supervision. (4) Operators can also open pop-up menus that provide specific information about the aircraft (see Appendix B, Image 7).

Each of the aircraft the participants monitored was identified on the screen with a specific symbol and call sign (e.g., DRAGON-13). The grid of targets and the region of control by the AWACS is illustrated in Appendix B, Image 6. By right clicking on an aircraft's symbol, it was possible to access specific information about the aircraft that was important for the memory items that followed (i.e., speed, heading, mission, altitude, and coordinate location). The target information for an aircraft is displayed in Appendix B, Image 7. Based on these seven attributes of the aircraft for a large number of targets (call sign, speed, heading, mission, altitude, location on grid map, and coordinate location), it was possible to generate a large number of unique memory test items (i.e., 114) for testing the hypotheses and deriving indices.

The participants were given instructions to monitor all of the aircraft in the area of responsibility. However, they were also told that they had primary responsibility over a subset of aircraft presented – the aircraft with the color corresponding to their weapons director icon on the screen display. Moreover, the participants were told that they were "responsible for monitoring and retaining knowledge concerning all of the attributes of each aircraft under your control." Consequently, the instructions highlighted the memory aspect of the task and that the participants were responsible for knowing all the characteristics of the aircraft under their own control as well as the aircraft that were under the control of the other weapons directors. These instructions did not vary by team size condition.

Training of Participants

The training of participants in performing the tasks associated with the weapons directors had a number of components. The participants initially received a general description of the synthetic task environment and were told about the different features of the workstation, screen, and controls. The screen display of a participant is presented in Appendix B, Image 6. After this, the specific features of weapons director's operations were described. This training was interactive allowing the participants to learn about different types of aircraft, their symbols, the functions the aircraft serves, and how the weapons director interacts with that type of aircraft.

After this initial training, the participants performed the weapons director tasks in a number of exercises and were given time to be come familiar with their workstations. During these exercises, the participants were quizzed to make sure that they understand the correct way to perform their functions. Problems, errors, and questions were addressed by the experimenter. Finally, a set of exercises were conducted to ensure that the participants achieved criterion levels of competence on the different tasks a weapons director performs in this synthetic task environment. Once the participant has reached these criterion levels, participation in the actual experiment began.

Procedure

After a general overview of the research project, the participants completed a battery of motivational (Motivational Trait Questionnaire, Heggestad & Kanfer, 2000) and personality measures (Big Five personality dimensions, International Personality Item Pool, 2001). Following these assessments, the participants continued with the training program until they reached criteria.

After a short break, the experimental session continued with a prebriefing regarding the situation they will confront. The participants were told that Turkey had suffered a devastating earthquake and requested assistance from the United States. The U.S. was providing food and water, medical assistance, construction equipment, and search and rescue teams. The

neighboring country to Turkey's west had shown hostility in the past, requiring patrol aircraft for that border. The aircraft under the AWACS's control would pursue five different missions: Air-drop food and water, combat air patrol, transport construction equipment, transport medical teams and supplies, and transport search and rescue teams. The cover story suggested that another AWACS was transferring control to the participants' AWACS because it was shifting off mission. During this exchange period, no critical events occurred so the participants would gain situational awareness in their task environment as well as focus on the attributes of the aircraft upon which they would be tested. The transfer of aircraft from the departing AWACS to the weapons directors occurred at a methodical pace so that the weapons directors had time to interact and respond to each of the aircraft operating in the scenario.

Forty five aircraft were present in the scenario at all times. In the five-person teams, nine aircraft were made the responsibility of each team member and were color coded with the responsible weapons director's color. For the three-person team, 15 aircraft were the assigned responsibility of each weapons director and were also similarly color coded. Participants were told to monitor all of the aircraft on the map, but that they had primary responsibility for the subset of aircraft that had an icon with their color. Participants were also told that they would be eligible to win a monetary bonus related to their ability to monitor and retain knowledge concerning the attributes of each aircraft during the simulation. Specifically, participants were told that they would complete examinations following the scenario and that the top 20% of performers on each of the examinations would receive a bonus of $10.

Once all of the aircraft were transferred to the weapons directors, the scenario continued relatively uneventfully for approximately 90 seconds. At this time in the scenario, the participants' display screens went blank. The participants were told that their AWACS had suffered an equipment failure. Because the weapons directors still needed to control their aircraft, it was important to know the situation they faced when screen information was lost. The participants were asked for their memory about the situation just before their screens went blank (first memory test). No discussion was allowed during this test period. The workstations were arranged with partitions between the weapons directors so that they would not be able to observe the actions and responses of the others. These partitions were pulled out before the participants began the scenario.

The recognition test involved 114 questions presented on the weapons directors' screens with responses gathered by pointing device clicks. Of these items, half were be true items (e.g., The altitude of the aircraft at this location was 2200 feet.) and half were false (e.g., The aircraft at this location was THUNDER 27.). This combination of items should establish a baseline with no bias toward a particular true or false response. The participants were asked to respond (clicking a button) indicating that the statement presented was true or false. In addition, the respondents were asked to indicate the confidence they have that their recognition response is correct on a 0-5 scale reflecting increments of 10% from 50% to 100%. The confidence scale also had anchors indicating 0 = 50-50, just guessing, 1 = slightly confident, 2 = somewhat confident, 3 = quite confident, 4 = very confident, 5 = completely confident.

After completing the first memory test, the participants in the individual condition completed a filler task. This filler task involved responding to a number of choice dilemma items (Kogan & Wallach, 1964). A set of 15 choice dilemmas were presented that asked the participants to determine what level of risk they would accept before they would recommend a person take a risky choice. The filler task took about 10 minutes to complete, so the time delay between the first and second memory tests was similar for the team and individual conditions. Once the filler task ended, the individual participants were asked to complete the second memory test.

Participants in the three- and five-person team conditions were also asked to respond to the second memory test, but as members of a weapons director team. The member of the team seated at the center was selected to enter the team responses for an item. The teams were told they could reach their team responses in any way they desired, but they were to make sure that their responses represent the collective opinion of all members of the team. After completing all of the second memory test items, the participants responded individually to a set of post-session questions concerning their beliefs about their performance on the test and interactions with their team members. After this questionnaire, the participants' questions were answered, they were debriefed about the study, information for distributing the incentives was gathered, thanked for their participation, and excused from the experimental session.

Video-records were made of the team interactions while performing the weapons director's tasks and of their collaboration during the second memory test. These video-records will be coded to determine if specific processes related to collaboration occur in the teams (e.g., error correction). This coding has not yet been conducted because it was not yet clear what of value should be coded from the interaction.

Results

General Memory Measures

The responses of the respondents in this experiment were analyzed in a number of similar ways. The recognition memory responses were aggregated to produce indices of memory (e.g., d', ∃, proportion correct, errors of omission, errors of commission). These indices were calculated for both the first and second memory tests. Mean values for these values are presented in the two panels of Table 8. An important general pattern of the data is that both individuals and teams achieved very low levels of performance on the recognition memory tests. These rates of memory performance were just above chance levels. This finding suggests that the memory load that weapons directors face may be very high. Weapons directors probably have and use memory aids that help them deal effectively with the large amounts of information that they must monitor when controlling a sample of aircraft.

It was predicted that the individuals and participants who responded as a team would not differ on any of these memory indices on the first memory test because random assignment to conditions would equate the conditions. Results were consistent with these expectations. There were no significant differences between conditions for the proportion of items answered correctly, $F(2,113)=1.45, p > .23$, sensitivity values of d', $F(2,113)=0.92, p > .40$, bias values of β, $F(2,113)=1.70, p > .18$, rate of errors of commission, $F(2,113)=0.59, p > .55$, and rate of errors of omission, $F(2,113)=1.55, p > .21$. The results also indicated that the

Table 8. *Mean Values and Standard Deviations of the Memory Measures for Recognition Test 1 and 2*

	Initial Recognition Test		
Dependent Variable	Individual M (SD)	Three Coactors M (SD)	Five Coactors M (SD)
Proportion Correct	0.51 (0.05)	0.52 (0.05)	0.50 (0.05)
d'	0.06 (0.26)	0.11 (0.26)	0.03 (0.33)
β (Beta)	0.98 (0.06)	0.98 (0.08)	1.05 (0.33)
Errors of Commission	.54 (.10)	.57 (.12)	.57 (.15)
Errors of Omission	.43 (.11)	.39 (.12)	.42 (.12)

indices did not differ between the first and second test for the individuals: proportion of items answered correctly, $F\ (1,25)=0.04, p > .84$, sensitivity values of d', $F\ (1,25)=0.12, p > .73$, or bias values of β, $F\ (1,25)=0.50, p > .48$, rate of errors of commission, $F\ (1,25)=0.64, p > .43$, and rate of errors of omission, $F\ (1,25)=0.25, p > .62$.

Second Recognition Test			
Dependent Variable	Individual M (SD)	Three-Person Team M (SD)	Five-Person Team M (SD)
Proportion Correct	0.51 (0.06)	0.55 (0.05)	0.51 (0.04)
d'	0.04 (0.30)	0.27 (0.27)	0.05 (0.22)
β (Beta)	1.00 (0.10)	0.95 (0.12)	0.99 (0.03)
Errors of Commission	.55 (.12)	.55 (.13)	.51 (.09)
Errors of Omission	.42 (.12)	.35 (.11)	.46 (.11)

The team responses on the second test provide values for the indices as benchmarks for normal levels of team memory performance on the AWACS synthetic task environment. Results indicated that the three-person teams outperformed the comparable individuals on the second test in measures of memory performance: proportion of items answered correctly, $F\ (2, 47) = 3.64, p < .05$, values of d', $F\ (2,47)=3.50, p < .05$. However, note that in Table 8, three-person teams were not significantly more biased (discrepant from $\beta=1.0$) than the individuals or five-person teams, $F\ (2, 47) = 1.41, p > .25$. Moreover, the five-person teams did not differ from the individuals or three-person teams on any of these measures.

Another way performance on the memory test can be examined is in terms of the errors observed. It was hypothesized that teams would have fewer errors than individuals due to the error correction process inherent in teams. Only somewhat supporting this hypothesis, the data indicated there were no differences in the rate in which individuals and teams were likely to produce errors of commission (false alarms/all false statements), $F\ (2,47)=0.48, p > .62$. However, three-person teams were less likely to produce errors of omission (misses/all true statements) than individuals and five-person teams, $F\ (2,47)=3.62, p < .05$. Given that three-person teams achieved higher levels of performance, this observed error correction probably contributed to the better performance.

These results suggest that five-member teams had lower levels on the measures of memory performance (e.g., proportion correct) than the three-member teams. Likewise, the five-person weapons director teams had more errors than three-person teams. These results are inconsistent with initial predictions. However, there was a relatively small sample of five-person teams for this analysis (nine). The results might become clearer and more stable with a larger sample of five-person teams.

Ideal Group d' Analyses

The responses to the recognition items provide a basis to test predictions from the ideal group model. The basic parameter estimates for the model are presented in Table 9. Calculations were conducted for both the three- and five-person weapons director teams separately. Although it was predicted that the five-person team might have better performance than the three-person teams, this was not found in the results reported above. Given the limited sample and relative poor performance of the five-person teams, it is obvious that the five-person teams did not even approach the optimal levels of performance as defined by the ideal group model. In fact, the results presented in Table 9 demonstrate that the five-person team did not improve on the initial individual levels of memory performance.

It is also of value to consider the performance of the three person teams. The results as presented in Table 9 do demonstrate that three-person teams did improve their performance above that of the initial member responses, but not by much. The three-person teams achieved only 23% efficiency, suggesting that they did not combine the available information nearly as effectively as predicted by the ideal group model. As more data become available and analyzed, it is hoped that better estimates of performance and parameters for the ideal group model will be forthcoming. Clearly, the results of this study with the AWACS teams do not agree with that from the memory studies reported above.

Table 9

Ideal Group Parameters and Analyses for Three- and Five-Person Teams

	Group Size	ICC Among Initial Responses	Mean d' of Team Members	Variance in Team Members d'	Ideal Group predicted d'	Observed Group d'	Efficiency (η)
Three-Person Team	3	0.351	0.110	0.065	0.568	0.272	23%
Five-Person Team	5	0.173	0.028	0.106	0.801	0.053	0%

Social Decision Scheme Analyses

Member Responses (C, I)	Team Correct	Response Incorrect
3,0	.71	.29
2,1	.61	.39
1,2	.50	.50

0,3	.35	.65

The responses of the teams on the second memory test were also compared to the different benchmarks established by social decision scheme theory. In particular, for the three-person teams, the single-correct, majority-correct, worst-case, and supraoptimal decisions schemes were considered. As with the evidence presented above regarding performance and the ideal group model, it is not reasonable to expect the teams to achieve optimal levels of performance as defined by the single-correct decision scheme. The observed distribution of responses for the different distributions of team member responses across the items are presented in Table 10.

Based on inspection of Table 10, it can be seen that teams exhibited a number of patterns. Much like the results presented in the memory studies presented above, as the number of team members who initially responded correctly increased, the teams were more likely to respond correctly. However, it was not as clear a pattern as that presented in Table 4. Also, for the three-person teams, there was a slight preference for the correct response in the team responses. This is ultimately reflected in the team performance being more accurate than individual memory performance. It is also important to note that in the 3-0 and 0-3 distributions, the unanimous agreement for the correct or incorrect responses did not mean the teams chose that alternative in their responses. For about 1/3rd of the items, the initially unanimous agreement resulted in a team choice for the other response. This pattern probably reflects the difficulty of the memory responses the participants had to make. The responses as illustrated in Table 10 and in general reflect the difficulty of making the correct response.

Table 10. Observed Distribution of Three-Person Team Responses Across All Recognition Items.

Model tests of the different plausible decision schemes against the observed responses did not provide clear support for any of the schemes. The single-correct, worst-case, and supraoptimal decision schemes were each clearly discrepant in their predictions related to the observed response, all $p <.0001$. The majority decision scheme best fit the responses, but it was still not a very suitable model, $\chi^2 = 6.71, p <.01$. A more extensive examination of other plausible decisions schemes not initially described above (e.g., proportionality, equiprobability) demonstrated that none of these decision schemes was a better predictor than the majority decision scheme. It should be noted that the majority-correct decision scheme is equivalent to a majority decision scheme for the three-person case because if a majority was not correct, then an incorrect majority was predicted to be likely to the dominate the teams' responses.

Social decision schemes analyses were also conducted on the responses from the five-person teams. The responses of these teams did not appear to differ from those of the individuals, so an improvement as a function of group judgment did not occur. This would suggest that a correct response was not more likely than incorrect judgment as was found in the studies reported above. The pattern of team responses based on the distinguishable distributions is presented in Table 11. Similar to the pattern of responses for the three-person teams, these teams were more likely to choose the correct response as the number of members initially choosing the correct response increased. Also, even when the team members were initially in unanimous agreement about the correct response (5-0, 0-5), the teams chose the other alternative about 1/3rd of the time. Moreover, in contrast to the plausible decision scheme described earlier for five-person teams, the pattern in Table 11 shows strong symmetry for the correct and incorrect responses for the comparable distributions.

Table 11. Observed Distribution of Five-Person Team Responses Across All Recognition Items.

Member Responses (C, I)	Team Correct	Response Incorrect
5,0	.67	.33
4,1	.66	.34
3,2	.56	.44
2,3	.44	.56
1,4	.39	.61
0,5	.36	.64

Model tests comparing the observed and predicted distribution of team responses for these five-person teams found that the majority decision scheme could not be rejected, $\chi^2 = .01, p >.92$, and that it provided a very good fit to the team responses. The other plausible decision schemes hypothesized earlier for the five-person teams (see Table 6; single-correct, pair-correct, strong-majority, worst-case, supraoptimal), could all be rejected as adequate models of the processes by which the responses were generated, $p <.0001$.

An inspection of the pattern of responses in Table 11 supports the acceptability of the majority decision scheme for the responses of the five person teams. However, because the initial distribution of member responses were relatively equivalent for the correct and incorrect responses (50% - 50%), similar predictions would be achieved by the proportionality and equiprobability models. So, although the majority decision scheme was an acceptable and the best fitting model from the decision schemes hypothesized earlier, it does not provide unique predictions relative to other decision schemes that summarize different underlying processes. However, given that the majority decision scheme was the better description of the team responses for both the three- and five-person teams, some basis for focusing on the majority process seems reasonable. This finding of a majority process of making team responses does reiterate the other results reported indicating that teams performed only somewhat better than individuals and that it was difficult for the participants to respond to the set of recognition items accurately.

Discussion

This experiment investigated team memory performance for information displayed to weapons directors in the AWACS synthetic task environment. It was predicted that collaboration in memory responses by weapons director teams would enhance performance over that of single weapons directors. This performance increment was observed for the three-person teams, but not the five-person teams. Research has consistently demonstrated teams are superior to individuals in cognitive task performance (Hinsz et al., 1997). This team superiority would be very useful for situations in which weapons directors lose the screen information in their displays. This experiment also tested benchmarks for memory performance in the DDD-AWACS task environment. The teams' performance when compared to the predictions of the ideal group model indicated the members did not approach optimal integration of the information they had available. Similarly, the social decision scheme analyses indicated that the teams were not performing near optimal. Rather, a majority process seems to describe the process by which the recognition responses were achieved in the teams. This study also provided data for constructing the β measure of response bias as well as the errors of omission and commission produced by teams and individuals. Consequently, from a number of perspectives, this study provided important evidence about the ways in which teams collaborate on a recognition memory task in an AWACS weapons director synthetic task environment.

Perhaps the most striking aspect of this study was the low levels of performance that were achieved on both the first and second memory tests by the individuals and teams. This poor performance is likely the result of the quantity of information that the participants were responsible for. With 45 aircraft having seven attributes each, a total of 315 pieces of information were presented. Moreover, the test asked for memory of the status of the aircraft at the time the screens went blank at the end of the scenario. Consequently, some of the information about the aircraft needed to be updated as the scenario changed over time (e.g., heading). This appears to have been a very challenging situation for participants to remember when they responded to the recognition task. Performance was barely above chance levels, and many participants may have been guessing for a large number of the items. This challenging information load actually was 1/3rd of what had originally been initially intended when the scenario was designed with the hope that a large number of recognition items could be used to establish stable parameters for the ideal group model. When the preliminary scenario was considered, it became obvious that the 135 aircraft were too many. Although the eventual scenario had only 45 aircraft, it is obvious that the load was still too challenging. As a consequence, a number of the hypotheses could not be adequately tested because the participants may have been responded by guessing because they had limited knowledge and memory for the information.

The patterns of the results for this study differed in important ways from those reported in the studies above. Performance in this study did not approach optimal levels. The social decision schemes that seemed to be most accurate in describing the team responses in this study did not appear to involve the correction processes of the earlier studies. And, the memory performance of the teams in this study was not clearly superior to that of the individuals. It is likely that the challenging nature of the memory items in this study contributed to the different patterns of results across studies. The team sizes in this study (three and five person) also differ from those of the earlier studies (two and six members). However, the patterns of results suggest that team size did not contribute to the differences.

An interesting implication of the differences in the patterns of results across the studies reported here is that the nature of the memory task itself influences the types of team processes that emerge. The items used in the group studies above appeared not to be as challenging although they involved rich social content. The ways the groups responded to that less

challenging task described above implied that near optimal performance involving error correction and information pooling were common to group performance on a recognition memory task. However, the pattern of results to this current study suggest that if the task becomes sufficiently difficult, these processes might change in important ways such that these mechanisms may not occur or are obscured by the difficulty of the items. As a consequence, the social decision scheme and ideal group model analyses produced results distinct from those of the earlier studies. Clearly the nature of the task, along with other features of teams, have direct and moderating influences on the ways teams perform cognitive tasks; a finding that this study demonstrated well.

This study was designed to be able to examine the performance differences resulting from different groups sizes. The models based on social decision scheme and signal detection theories can be applied to determine the implications of changes in the size of weapons director teams for memory performance. The predictions of differences for three and five member teams were made; however, the results did not support these predictions. In fact, the results reflect an unexpected decrease in performance for five-person teams compared to three-person teams. More responses from three and five member teams will be required before this difference can be held with confidence. So, at this time, no conclusions about the impact of reducing AWACS teams from five to three members can be made. With data gathered from more teams, additional issues can be addressed as well. For example, the confidence measures included in this study might help illuminate the nature of the difficulties that the participants had in responding to the set of recognition items. In addition, the motivation measures assessed during the first session can also be combined with the general performance measures to help us understand what makes teams and their members effective as weapons directors.

This study added to the literature that examines how the contributions members of a team are combined to produce the team outcomes and performance. Social decision scheme theory provides one approach to this issue. The observed decision schemes presented in Tables 10 and 11 revealed a number of characteristics of weapons director team memory performance. The observed decision schemes were inspected to discern the processes that best summarized the team memory responses. This inspection was independent of the more objective model testing approach. A majority process appeared to account for the responses of these teams. A majority process is consistent with other group decision making literature; however, it differs from that reported in other research on collective memory (Hartwick et al., 1982; Hinsz, 1990; Hinsz et al., 2005; Vollrath et al., 1989). A majority process often results on tasks that have no correct answer and the members are deciding on issues of opinion. The recognition items used in this study did have a correct response, however, given the low probability of answering correctly, perhaps the members could not distinguish a response as being factually correct. If it was not possible to demonstrate that one response was the correct answer (Laughlin, 1980), then it would be reasonable that the teams would treat the items as matters of opinion, and select a team response based on a majority process.

One feature of this experiment is that the weapons directors will have responsibility for specific aircraft matching the color of their icon. Consequently, the different weapons directors should have had specific knowledge or expertise with regard to the different aircraft on the display. This differential knowledge could lead teams to allow member expertise to play a role in how information is remembered. It could happen that the weapons director with controlling responsibility for an aircraft would have greater influence on the team responses for information about that aircraft. This suggests an expertise-based decision scheme: if the weapons director with responsibility for the aircraft remembers a piece of information for the aircraft, then the team accepts that piece of information as correct (cf., Kirchler & Davis, 1986). This expertise-

based decision scheme would be similar to a single-correct decision scheme, and so would be more optimal than the pair-correct decision scheme. Further analyses of the team responses with a larger sample will be conducted to consider the role that expertise might have on effective team performance.

The memory performance of the weapons director teams in this study contributes to the body of research that uses the DDD-AWACS synthetic task environment. These data provide insights for weapons director team effectiveness which can be combined with related research efforts. This study also used theoretical models of team performance to establish benchmarks for optimal memory performance. Although the ideal team *d'* and supraoptimal decision scheme models provided ways to identify extraordinary performance if it arose, there was no evidence that it occurred for these teams. Given the poor performance of the individuals and teams in performing the memory task, this should not be expected. Nevertheless, this does raise the question about what about the recognition items might have contributed to the lower than expected levels of performance. The information that was provided could be perceived as challenging and the amount of material to remember might overload normal memory capacity. But it might also be the case that the type of material used in this experiment -- the information provided on the AWACS weapons director's display -- could also contribute to the poorer performance. Is certain information that weapons directors receive more difficult to remember than other information? This is the issue raised in the next section which has implication for potential modifications in how information should be displayed to weapons directors.

11. <u>Differential Memory for Weapons Directors' Display Information</u>

The simulated AWACS task environment described in experiment above provides data that is relevant to issues regarding memory for different types of displayed information. One aim of the previous experiment was to determine if specific attributes that are displayed to the weapons directors are remembered better than other types of information. For example, will the location of an aircraft be remembered better if it is displayed on the map grid than if it is given as coordinates? An important human factors question is if certain types of information (e.g., semantic – mission) is more likely to be remembered correctly than other information (e.g., numerical – altitude). If different types of attribute information are more likely to be remembered accurately, then it may be appropriate to consider modifications that will assist the memory of critical information that is remembered less well. There are strong theoretical reasons to believe that different kinds of attribute information displayed to weapons directors will be remembered more or less well (Hinsz & Malone, 2004). The experiment described directly above provides an opportunity to test these theoretical implications with information displayed to simulated weapons directors in an AWACS synthetic task environment.

Much of the following theoretical discussion of memory for different types of information is based upon a published chapter [Hinsz, V.B., & Malone, C.P. (2004). Memory for attributes of information presented in a synthetic task environment: An illustration with AWACS weapons directors' displays. In S.G. Schiflett, L.R., Elliott, E. Salas, & M.D. Coovert (Eds.), *Scaled Worlds: Development, validation, and applications* (pp. 220-239). Hampshire, England: Ashgate Publishing Limited.] However, the empirical analysis of the data from the experiment has not yet been presented elsewhere.

It is important to understand how the simulated AWACS weapons director task environment interfaces with the ways humans normally process information. If the information is presented in a fashion that conflicts with the ways humans process the information, then weapons directors might not perform at levels of normal benchmarks. However, if the

information is presented in a way that corresponds to the way it will be optimally processed, then weapons directors should be able to demonstrate task performance at the levels of their capabilities.

This study and analysis of data from the experiment just presented addresses questions regarding how the way information is presented and displayed may interact with cognitive processes associated with human memory. For many task environments that team members in the Air Force perform, higher-level cognitive processes of inference, judgment, reasoning, and problem solving are involved. These higher-level cognitive processes often rely heavily upon memory processes (Hinsz, 1990). Being able to remember critical information is important for the inferences and responses that follow. If the information was not stored initially, or if the information was encoded in a way that makes it difficult to retrieve, or if retrieval fails, then that critical information cannot be brought to bear on the issue being addressed. Consequently, it is important to understand how the way information is displayed influences weapons directors' memory for that information.

Human memory is a system for storing and retrieving information for subsequent analysis. In order to systematically examine the memory processes that may be involved in the weapon directors' synthetic task environment, the way in which visual stimuli are processed and remembered is described, highlighting inherent functions likely to affect performance. This discussion focuses on the two active stages of memory processes—encoding and retrieval. Encoding refers to the process of acquiring information and forming a memory record. Retrieval involves to processes of accessing information from long term memory so that it can be used for some objective. Both the encoding and retrieval processes are discussed for their implications for memory of different kinds of information from the screen displays of weapons directors.

Memory systems are limited with respect to the amount of information acquired and stored after a very brief exposure. Approximately eight items of visual information are completely available in the form of a snapshot of the scene, if accessed within 1 second of exposure (Sperling, 1960). The basic function of this brief representation is to prolong very brief stimuli and guarantee them a minimal persistence, thereby increasing the chance that the stimuli will be present to be analyzed further (Efron, 1970; Haber & Nathanson, 1968). The limitation on the duration of this early memory seems to be accommodated by the AWACS weapons director environment, as visual information is presented for approximately 12 seconds before updating (i.e., sweep rate on the display). However, in the AWACS synthetic task environment, a potentially large number of elements can be presented in the graphic display at one time, requiring more than one fixation to perceive all elements. In addition, a second display embedded in the graphic display provides detailed information regarding speed, altitude, name, heading, mission, and coordinates (see Appendix B, Image 7). Display of this information is controlled by the operator using pop-up windows; therefore, the likelihood that the operator attends to the information upon presentation increases. However, revealing the embedded information requires effort. The information in the pop-up menus can not be obtained in a cursory glance as with the graphically displayed information.

After initial processing of information, the relevant information may be retained briefly in immediate memory if it is needed for further processing. Cognitive work or processing takes place in immediate memory (also called short term memory or working memory). Because an operator must evaluate incoming information and use the information to make decisions for future action, immediate memory is particularly important in the AWACS environment. Information in immediate memory decays within 20 seconds unless the information is rehearsed in some way (Peterson & Peterson, 1959). Repetition is an effective maintenance strategy

involving repetition of the visual information to oneself (usually silently). Repeating visual information to oneself is thought to be a process of re-coding the visual features of the stimulus into acoustic features. After all, sounds are more amenable to quick rehearsal than are visual components of letters or numerals. In the AWACS environment, operators may verbally summarize and rehearse key information in order to maintain the trace in immediate memory. For example, "THUNDERs 2 through 4 heading 235 at 10,000 feet."

Capacity is another aspect of immediate memory relevant to memory performance in simulated environments. After a single presentation, individuals typically recall an average of seven units of information (Miller, 1956). Individuals may group (or chunk) separate pieces of information into higher-level units in order to retain more information in immediate memory. The process of chunking involves using meaning to impose order on individual stimuli in the form of higher-level units, thereby increasing the raw capacity of immediate memory. A related process occurs when groups of friendly aircraft fly so closely that they appear as a single track on the display. The operator, in effect, is able to encode this group of aircraft as a single unit (i.e., a flight), provided he or she also encodes how many individual aircraft compose the unit. This grouping strategy should occupy less space in working memory, leaving room for other storage tasks.

Although there certainly are capacity constraints in immediate memory, researchers have demonstrated that while engaging in a standard digit span task (briefly storing a string of digits for immediate recall), humans perform quite well on a second task involving a true/false reasoning task (Baddeley, 1986). Studies like these provide evidence for the involvement of various cognitive processes such as attention and reasoning in immediate memory, hence the notion of working memory. Therefore, operators should be able to store status information about individual targets in the graphical display (e.g., relative locations of friendly and enemy aircraft) while making decisions and reasoning about required actions indicated by the pop-up menu displays (e.g., is a fighter nearby).

In the AWACS synthetic task environment, the quantity of information is too great to be stored and processed by one weapons director. Accordingly, weapons directors work in teams in which they are each primarily responsible for a subset of targets. Targets are color-coded and indicate which weapons director is responsible for each subset. The weapons director in charge of blue targets, for example, will perform optimally when his or her attention is focused only on blue targets, especially during busy points in the mission. This division of labor accommodates the limited storage and processing capabilities of immediate memory.

According to one popular model incorporating storage and processing capabilities in immediate memory, working memory is composed of three main systems operating together (Baddeley, 1992). The processing component or central executive controls and monitors a person's attention according to needs current at a given time. This central executive is particularly important in activities that require one to attend to and manipulate large amounts information in the presence of time constraints. For operators in the DDD environment, central executive functions are particularly important in shifting attention between aircraft in order to service them in a timely manner. Two storage systems operate under the control of the central executive. The visuo-spatial sketch pad (Baddeley, 1992) stores and manipulates visual images (e.g., the spatial arrangement of aircraft) and the phonological loop stores and rehearses speech-based information (e.g., identifying information for each aircraft).

Although the functions of the central executive allow for the execution and manipulation of relatively large amounts of information over a short period of time, processing resources are limited in immediate memory. For example, Teasdale and colleagues (1995) demonstrated than an irrelevant activity such as daydreaming degrades performance on a task of generating a

continuous string of random digits—a task likely directed by the central executive. This implies that tracking and commanding a target will be impaired by an additional cognitive activity, such as reading a text message in the communication text box or changing control settings in the AWACS environment.

In general, pictures are better remembered than words describing the pictures; a picture-superiority effect (Madigan, 1983). Therefore, information in the weapons director's with visual-spatial representations should be remembered better than other attributes. Pictures are particularly well-suited for presenting certain types of information, but not others. Pictures are very effective in presenting concrete and spatial information, as this information is displayed in a manner analogous to the object or event being represented. Abstract and categorical information, however, are not amenable to visual representation and are better conveyed by verbal representations. For example, in the AWACS synthetic task environment graphical display, relative locations of aircraft are amenable to pictorial representation in memory, while each aircraft's mission may be better represented with a verbal representation.

The type of processing occurring at encoding has been demonstrated to result in differential levels of memory performance (Craik & Lockhart, 1972). If the objective is to store information longer than a few minutes, deeper levels of processing, such as focusing on the meanings of the to-be-remembered items, will result in better long-term memory performance than shallow levels of processing, such as rote repetition. Encoding tasks that require analysis of the entire word encourage elaboration, which involves processing information deeply by assimilating it to one's prior knowledge through mnemonics, imagery, or organizational schemes. Operators will display better memory for presented information if they acknowledge receipt of the new information and incorporate it into existing visual or verbal representations.

There are a number of implications of several retrieval factors for memory performance in the weapons directors' AWACS simulation environment. Retrieval is generally defined as the process of accessing and using information stored in long term memory. The relationship between encoding and retrieval contexts affects the degree to which retrieval is successful. In general, retrieval is successful to the extent that relevant cues are present during retrieval (Tulving, 1974). That is, retention is best when the internal and external retrieval cues match the cues used during study. A retrieval cue acts as a reminder when accessing a particular piece of information. Memory performance will be best if that aircraft label is provided as a cue when retrieval is required. The weapons directors' memory performance should be better if asked about the aircraft THUNDER 25 than if asked about the aircraft whose altitude is 1200 feet. Note that altitude is likely to be a redundant feature across aircraft and is not likely to be a distinguishing characteristic upon encoding. Consequently, a prediction is that questions that referred to by its identification information will be more likely to be remembered than if it is based on its location on the map grid.

Retrieval is a distinct and crucial phase of the memory process. It involves accessing stored information so that it can be used with other information in making inferences or decisions. Generally, the ease of accessing the desired information is associated with the conditions present when the information was initially understood and encoded. If the learning conditions can be sufficiently replicated at retrieval, information can be stored for indefinite intervals and successfully brought out of storage and used for the task at hand. The encoding aspects of memory, however, tend to reflect more the information that is held for shorter durations. Regardless, this consideration of encoding and retrieval has provided a good foundation for understanding how the way information is presented in the AWACS environment might influence the memory of operators.

There is an assortment of implications of human information processing that are relevant

to the consideration of memory for information in an AWACS workstation display. For this discussion, the implications are organized around encoding and retrieval processes. The analysis of the encoding stage of memory points to several important considerations.

Differential levels of memory performance result for pictorial and text forms of presentation. The picture superiority effect (Madigan, 1983) suggests an interesting hypothesis for information displayed in the AWACS task environment. Information displayed in pop-menu format (e.g., speed, heading, altitude) in numeric format should be remembered less well than the visio-spatial information from the grid such as the location of the aircraft or target.

Deep levels of processing generally lead to higher levels of performance than shallow levels of processing (Craik & Lockhart, 1972). Deeper levels of processing focus on the meaning of the to-be-remembered items, as well as relationships between those items. This finding suggests that the semantic information associated with an aircraft should be remembered better than information that has less meaning. Consequently, information about an aircraft's mission should be remembered better than information about its heading which would require more effort to process in a fashion to give it strong meaning.

The analysis of the retrieval stage of memory has also pointed to several important considerations in memory for information in the AWACS simulation task. Important consideration should be given to degree of match between cues present during study and testing situations and salience of cues during study.

Memory performance is determined, in part, by the degree of match between cues present during encoding and later use (Tulving, 1974). Some aspects of forgetting can be explained as the lack of effective cues to enable retrieval. Information unable to be recalled may be successfully retrieved with appropriate hints or reminders. This implies an advantage for consistently displaying key information throughout a mission or session (e.g., target label, heading, direction, altitude, etc.). All information present at the time of learning the target information has the potential of being an effective retrieval cue for that piece of target information. Consistency in the types of information displayed will assure a greater degree of match between cues during learning and later task performance. Because of these findings, the recognition test used consistent formats in displaying the information that was also used in referring to the aircraft in the recognition item statements. About half the aircraft were referred to in terms of the identification information (THUNDER 38) while the rest were referred to by the location on the map grid (i.e., the aircraft at this location).

This analysis of encoding and retrieval of information displayed to weapons directors has led to a few hypotheses and highlight some other important considerations for testing the memory for this information. Although the primary aim is on seeing which types of information receive the more accurate memory responses, it is also of interest to examine the nature of errors for the different kinds of information. That is, are the types of errors that emerge influenced by the types of information (e.g., alphanumeric) being considered. Is semantic information more prone errors of omission than other information (e.g., numeric), which is more prone to errors of commission. These issues as well as others discussed above are investigated with the data gathered as part of the experiment described directly above.

Method

The experimental method for this analysis is based on that of the experiment described directly above. The same 116 participants were involved with analyses based on responses to the 114 recognition items on the first memory test. During the first recognition memory test, all participants were responding as individuals and had not received any instructions about working as a team. Consequently, the data that were analyzed were responses from participants to the

recognition test items on the first test.

The recognition items were divided into a number of distinct categories reflecting the type of information that was displayed. The recognition items always referred to an aircraft that was displayed during the scenario in one of two ways. For 54 of the recognition items, the aircraft was referred to by name, e.g., "DRAGON 35 was at a heading of 349 degrees." For the other 60 recognition items, the aircraft was referred to by a red crosshair (see Image 8, Appendix B) that was placed on the map grid, e.g., "The aircraft at this location was at an altitude of 26000 feet." Consequently, all recognition items referred to the aircraft by name or location.

The recognition items were further categorized by the type of information that was assessed. For the items referring to the aircraft by name, the (a) speed, (b) heading, (c) altitude, (d) mission, or (e) coordinates were the attributes the participants' had to judge as being stated accurately (true) or inaccurately (false). For the items that referred to the aircraft by location, most of the same information would be the critical information to be judged as accurate or inaccurate (i.e., speed, heading, altitude, mission). However, for items based on location, 16 of the items referred to (f) the name of the aircraft (e.g., The aircraft at this location is DRAGON 4.). Thus, all recognition items are divided into statements that refer to an aircraft by name or location, and further categorized into one of six kinds of attributes. The numbers of items reflecting these categories are presented in Table 12.

Table 12. The Number of Recognition Items as a Function of Attribute Class.

Attribute Assessed	Aircraft Identified by Name	Aircraft Identified by Location
Speed	10	10
Heading	10	10
Altitude	10	10
Mission	14	14
Coordinates	10	-------
Name	-------	16

As indicated for the experiment above, all participants responded to all 114 recognition items in the first test individually. The participants responded that the statement given was true or false. Then the participants indicated how confident they were in the accuracy of their true-false response. For the responses to the recognition items in the first test, the participants were offered a $10 bonus if they were among the top 20% of performance in terms of the number of correct responses. It is also important to note that participants were admonished to do well on the test and to answer as best as possible with correct response for each statement.

Results

Analyses of the responses were primarily conducted based on a repeated measures analysis of variance. Because each individual responded to each of the classes of aircraft attributes, comparisons will be made across participants for the different attributes. The mean proportion of recognition items answered correctly was the data used for each participant. The analysis indicated that memory performance varied significantly as a function of attribute class, $F\,(5, 570) = 9.02, p < .0001$. Mean proportion correct responses by attribute class are presented in Table 13. Post-hoc comparisons using the Tukey B test indicated that mean values for each attribute class were significantly different from every other attribute class. There was no effect of eventual team condition (individual, three person, five person) on proportion of items answered correctly, $F\,(2,112) = 1.46, p > .23$.

It is also possible to examine the proportion correct response as a function of whether the aircraft was identified by name or location. An inspection of Table 12 indicates that questions regarding aircraft speed, heading, altitude, and mission were each assessed with aircraft name and location used to identify the aircraft. Therefore, a 2(Identifier: Name or Location) x 4(Attribute: speed, heading, altitude, and mission) repeated measures analysis of variance can be informative regarding the relative impact of name or location identification on memory performance. Mean proportion correct values for these conditions are also provided in Table 13. There was a significant influence of name ($M = .531$) versus location ($M = .488$) on the proportion of these types of items answered correctly, $F(1, 114) = 15.93, p < .0001$. There was also an identifier by attribute interaction for this analysis, $F(3, 342) = 6.07, p < .0005$. Further examination of Table 13 indicates that responses were more often correct for speed, altitude and mission for aircraft identified by name than by location, but that this difference didn't appear to occur for heading. It is also interesting to note from Table 13 that it was easier to answer items about an aircraft's name given its location on the grid than items about the aircraft's coordinate location given its name. All of these results suggest that the type of information being encoded and retrieved had a strong influence on the accuracy of the memory response.

Table 13. Proportion of Recognition Items Answered Correctly as a Function of Attribute Class.

Attribute Assessed	Aircraft Identified by Name	Aircraft Identified by Location	Overall
Speed	.555 (.18)	.481 (.13)	.518 (.11)
Heading	.503 (.13)	.506 (.15)	.504 (.09)
Altitude	.504 (.17)	.488 (.15)	.496 (.12)
Mission	.562 (.13)	.477 (.12)	.520 (.09)
Coordinates	.447 (.16)	-------	.447 (.16)
Name	-------	.544 (.13)	.544 (.13)

Discussion

In this study, the AWACS weapons directors' synthetic task environment was used to address the question of how well simulated operators could remember the types of information displayed. This study concerns the issue of how the way information is displayed to operators interacts with basic human cognitive processes to influence the operators' capability to remember the information that was displayed. This study was not an exhaustive analysis of the problem. However, this examination summarizes some of the issues as well as provides some implications for considering changes in the way information should be displayed as well as the need for memory aids. This study provided evidence about the rates of remembering the different types of information displayed on the weapons directors' screen reflecting graphical, spatial, semantic, symbolic, and numerical information.

The results of this study supported a number of the hypotheses stated earlier. In particular, the recognition items that referred to the aircraft by name were more often answered correctly than items that identified the aircraft by its location. Also, attribute information having a numerical value (e.g., heading, altitude) was remembered less well than attribute information of a semantic nature (mission) or the symbolic identity of the aircraft (name). It is of interest to note that speed information was also recognized well. This finding implies that there is something about aircraft speed that made it more memorable. For example, perhaps the participants could cognitively represent aircraft speed in a fashion of physical speed through the

display grid. It would be necessary to assess how speed, altitude, and heading information are represented to determine why some of this numerical information was recognized more accurately than for other attributes. The results of this study do demonstrate that the responses of the participants provide clear evidence in support of the hypotheses that differences in memory would arise for the different attributes of the aircraft.

It is clear that the theory and research on memory processes can inform us much about the ways that operators would remember material that is presented in the AWACS synthetic task environment. This study also shows the interplay between theory and application. This examination of the way information is displayed in the weapons directors' task environment raises new questions to be addressed by theory and research. In particular, it was intriguing to discover that static identification information (the aircraft's mission) was remembered better than dynamic information about the aircraft's last location. This pattern of responses provides evidence for the multiple memory systems view (Schacter & Tulving, 1994). These results suggest that separate memory systems exist for storing particular types of information (e.g., spatial and semantic) and that memory performance is affected by tasks that require one particular system compared to tasks that require activity in more than one memory system. It is also the case that the patterns of correct responses favoring specific attributes suggest that the design of the weapons directors displays might be altered in order to optimize performance. The applied situation as reflected in this study suggests that weapons directors have circumstances that are information rich and high in cognitive demands, which makes this study a contribution to theoretical accounts of memory, attention, and knowledge representation.

12. Summary and Conclusions

This report describes a program of research investigating how collaboration influences collective performance on recognition memory tasks. The studies were conducted with a two different types of information to be remembered. Moreover, the studies compared the memory performance of two, three, five, and six member groups with those of individuals treated similarly. Memory performance of these groups and individuals was examined for measures of correct responses and errors, as well as other features of memory. Also, the memory performance for different kinds of attributes of the display of aircraft in an AWACS environment was investigated. Across these studies described, theoretical models of team performance were applied and tested. These models provided a basis for understanding the levels of performance observed relative to benchmarks for optimal and suboptimal levels of performance.

There are an intriguing set of patterns for the results for the studies reported. As expected, groups and teams generally outperformed comparable individuals. A portion of this difference can be attributed to groups having fewer errors, in particular errors of omission. However, the pattern of results appeared to be influenced by group size. Groups of six members seemed to achieve higher levels of performance relative to dyads. Additionally, when the material was very difficult to remember, the performance of five-person teams did not differ from that of individuals, although three-person teams were better at recognizing the information than individuals. So, the results of the studies also suggest that the difficulty of remembering the material presented influenced the processes by which the teams performed recognition memory tasks. The final study also showed that the nature of the material influenced how it would be remembered. Consequently, the studies reported illustrated many results that contribute to knowledge about how teams and individuals perform tasks involving recognition memory. In the process, these studies also contribute to the research on team judgment and individual cognition.

Each of the studies reported also make contributions to our understanding of collaboration in performance of cognitive tasks. The first study in this report indicated that performance beyond that predicted by the ideal group model might be possible with memory performance. A second study attempted to replicate this finding with smaller groups but revealed less support for extraordinary levels of performance. The second study also demonstrated that there was no support for motivational factors being the basis for the extraordinary performance reported in the first study. Social decision scheme analyses were also reported for the group responses of the first and second studies. Both studies indicated that groups were more likely to select the correct response as the number of initially correct members increased. The initial two studies set the stage for the next two studies that examined performance of simulated AWACS weapons directors. A key finding of the last two studies was that the cognitive situation the simulated weapons directors faced had very challenging cognitive demands. Participants' memory performance was generally slightly above chance levels. As a consequence, the three and five member teams in the third study did not show performance approaching optimal levels, but rather performance was quite suboptimal. A social decision scheme analysis suggested that a majority process best accounted for the team responses in the third study. This majority process is likely the reflection of the difficulty of the team knowing if they could determine the correct response. The fourth study extended this analysis of the impact of the characteristics of the information displayed to the simulated weapons directors. Consistent with hypotheses, attribute information that was numerical in fashion was more difficult to remember correctly compared to semantic, spatial, and symbolic information. Also, the simulated weapons directors responded correctly more often if the aircraft being considered was referred to by its name than by its grid location. Therefore, as a consequence of this series of studies, a number of important questions about collaborative performance were addressed, but several other questions were raised.

An aim of the research reported was to provide a foundation for considering ways of intervening to improve performance of weapons director teams. One issue considered was the potential implications of reductions in weapons directors' team size. The pattern of results across the studies suggest that there a number of potential negative consequences of reducing team size. The cognitive demands of the task seem to be very challenging and sharing responsibility for aircraft would be one way to reduce this cognitive load. The predictions from the theoretical models clearly indicate that performance would suffer if team size was reduced. The results of these studies suggest that not only will performance decline, but the team processes that lead to the performance would be less effective as well.

Another aim of this research program was to explore benchmarks for performance of simulated weapons director teams. The initial study indicated that groups achieved optimal levels of information integration when responding to the recognition items. The second study did not achieve this level, but the dyads' performance still showed substantial improvements in performance compared to individuals. The study involving the AWACS weapons director task indicated that teams were quite suboptimal; the five-person teams were not better than the average of individuals. Model testing with social decision schemes indicated that a majority process accounted for the responses, which was suboptimal for the task considered. In contrast to the initial study, none of the following studies replicated the finding of a potential supraoptimal performance of the teams.

Results from these studies suggest that one way a team might improve its level of performance is by relying on the expertise of the weapons directors for their own areas of responsibility. This reliance on expertise resembles transactive memory processes considered to contribute to effective team performance (Wegner, 1987, 1995; Hollingshead, 1998; Moreland, 1999). Future experiments in this research program aim to assess the influence of transactive

memory processes in weapons director teams. Such research is particularly compelling because weapons directors can follow the functional assignment of responsibilities or a geographical assignment of responsibility in which a weapons director is responsible only for aircraft in their portion of the screen display. It may be that a functional assignment of duties encourages the development of a transactive memory system that enhances memory performance, whereas a geographically-based assignment might not result in a transactive memory system and memory performance suffers as a consequence. Future experiments could address questions of this sort that consider how teams develop effective strategies for performance.

This set of experiments contributes an understanding of how the collaborative memory processes of weapons director teams can lead to enhanced effectiveness and superior performance of these teams. This research helps provide a basis for understanding how team performance of weapons directors can be superior to that of individuals. The research reported used a synthetic task environment to investigate the performance of weapons director teams. By conducting the research using this synthetic task environment the results can be shared and compared with the work of other researchers using the AWACS weapons director task. Moreover, this research has potential implications for other command and control teams structures for which collaboration can be used to harness the advantages of teams.

The research reported from this funded project is related to a number of other research topics. While conducting the research reported above, the principal investigator was also involved in a number of other research projects. Several of these projects are related to the topics relevant to the main objectives of the research reported. These projects are related to group decision making, group and team performance, group and individual goal setting, and judgment and decision making. Many of these projects were pursued with the assistance of graduate students and post-doctoral scholars; many who were supported in part with funds from this grant. In addition to the topics related to the preceding research, research was also conducted on other projects dealing with social-organizational psychology phenomena. These various research projects benefited from the interaction of ideas and methods from these related projects. Consequently, the next focus of this report is brief discussions of the supporting research efforts that resulted from the period of grant funding.

13. Supporting Efforts in Group Decision Making

Much of the research reported in the studies reported above have their bases in the research on decision making in groups and teams. Social decision scheme theory has had its greatest impact on research on group decision making. Much of the research on memory performance in groups also has its basis in group decision making research (Hinsz, 1990). Consequently, other research conducted by the research team related to group decision making was often of direct relevance to this grant supported effort. Below the publications and manuscripts related to group decision making completed during the period of the grant are listed.

Tindale, R.S., Kameda, T., & Hinsz, V.B. (2003). Group decision making. In M.A. Hogg & J. Cooper (Eds.), *Sage Handbook of Social Psychology* (pp. 381-403). London: Sage.

Hinsz, V.B. (2004). Metacognition and mental models in groups: An illustration with metamemory of group recognition memory. In E. Salas & S.M. Fiore (Eds.), *Team Cognition: Understanding the Factors That Drive Process and Performance* (pp. 33-58). Washington, DC: American Psychological Association.

Hinsz, V.B., Henkel, J.M., & Tindale, R.S. (in press). American military courts martial: Processes and procedures of trials and decisions. In M.F. Kaplan & A.M. Martin (Eds.), *Understanding world juries through psychological research.* New York: Psychology Press.

Hinsz, V.B., Tindale, R.S., & Nagao, D.H. (in press). The accentuation of information processes and biases in group judgments integrating base-rate and case-specific information. *Journal of Experimental Social Psychology.*

Hinsz, V.B. (submitted). Group judgments of the frequency of events: Accuracy, bias, knowledge transfer, and out-of-range responses. *Group Processes and Intergroup Relations.*

Lawrence, D.M., & Hinsz, V.B. (submitted). Group members as actors and observers in attributions of responsibility. *Group Dynamics.*

14. Supporting Efforts in Group and Team Performance

A main objective of the grant funded research effort was to understand the factors that contribute to team performance. The ideal group model and social decision scheme approaches allowed the development of benchmarks for effective and ineffective performance on the memory task. Moreover, the group and team performance was assessed on a cognitive task. The research on group and team performance served as an important conceptual basis for the research conducted with the grant funding. Therefore, it is appropriate to consider other research on group and team performance that the researchers pursued during the period of the grant. Below publications and manuscripts related to group and team performance are listed.

Meier, B.P., & Hinsz, V.B. (2004). A comparison of human aggression committed by groups and individuals: An interindividual-intergroup discontinuity. *Journal of Experimental Social Psychology, 40,* 551-559.

Park, E.S., Hinsz, V.B., & Ladbury, J.L. (in press). Enhancing coordination and collaboration in remotely operated vehicle (ROV) teams. In N.J. Cooke, H. Pringle, H. Pederson, & O. Connor (Eds.), *Human Factors of Remotely Operated Vehicles*. North Holland: Elsevier.

Reimer, T., Park, E.S., & Hinsz, V.B. (in press). Shared and coordinated cognition in competitive and dynamic task environments: An information-processing perspective for team sports. *International Journal of Exercise and Sport Psychology.*

Park, E.S., & Hinsz, V.B. (in press). "Strength and safety in numbers": A theoretical perspective of group influences on approach and avoidance motivation. *Motivation and Emotion.*

Hinsz, V.B. (revision requested). Meeting and exceeding expectations of ideal group performance: Information integration in group memory performance. *Journal of Experimental Social Psycholoigy.*

Park, E.S., & Hinsz, V.B. (submitted). Group interaction sustains positive moods and diminishes negative moods. *Psychological Science.*

15. **Supporting Efforts in Group and Individual Goal Setting**

The principal investigator for this grant has an established research program on group and individual goal setting. This research reflects the principal investigator's interests in decision making and performance. Research in goal setting provides a unique opportunity to investigate decision making and performance in one setting. In this research, individuals or groups are asked to set a goal for their performance on some task. Goal setting research routinely finds that if the proper goals are established, these goals can enhance performance on the task. Consequently, after establishing task-performance goals, the individuals and groups are asked to perform the task for which they established a goal. In this way, the goal decision has direct implications for the task performance that the individuals and groups pursue. Research completed during the period of this grant examined a number of factors related to the setting of goals, the relationship between goals and performance, and task performance. The manuscripts and publications related to goal setting by groups and individuals are listed directly below.

Spieker, C.J., & Hinsz, V.B. (2004). Repeated success and failure influences on self-efficacy and personal goals. *Social Behavior and Personality, 32*, 191-198.

Henkel, J.M., & Hinsz, V.B. (2004). Success and failure in goal attainment as a mood induction procedure. *Social Behavior and Personality, 32,* 715-722.

Hinsz, V.B., & Nickell, G.S. (2004). Positive reactions to working in groups in a study of group and individual goal decision making. *Group Dynamics, 8,* 253-264.

Hinsz, V.B., & Jundt, D.K. (2005). Exploring individual differences in a goal-setting situation using the Motivational Trait Questionnaire. *Journal of Applied Social Psychology, 35,* 551-571.

Hinsz, V.B. (in press). The influences of social aspects of competition in goal-setting situations. *Current Psychology.*

Hinsz, V.B. (revision requested). Competitiveness and competition influences in goal-setting situations. *Human Performance.*

Heimerdinger, S.R., & Hinsz, V.B. (revision requested). Failure avoidance motivation in a goal-setting situation. *Human Performance.*

16. Supporting Efforts in other Judgment and Decision Making topics

The research reported above involved memory judgments. These judgments were made by individuals and groups about material they were or were not presented. Consequently, research on judgment and decision making is related to this research. Signal detection theory and social decision scheme theory both have their bases in the research on judgment and decision making and contribute to that research as well. Given the relevance of judgment and decision making research to the research reported above, supporting research was also completed related to the judgment and decision making field. Below are listed a number of publications that contribute to the individual judgment and decision making literature completed during the period of this grant.

Jundt, D.K., & Hinsz, V.B. (2002). Influences of positive and negative affect on decisions involving judgmental biases. *Social Behavior and Personality, 30*, 45-52.

Magnan, R.E., & Hinsz, V.B. (2005). Mood, gender, and situational influences on risk-taking advice for others. *Social Behavior and Personality, 33*, 1-10.

Hinsz, V.B., Heimerdinger, S.R., Henkel, J.M., & Spieker, C.J. (2005). Test-accuracy and base-rate information in the prediction of disease occurrence. *Current Psychology, 24*, 80-90.

17. Additional Efforts in other Social-Organizational Psychology topics

During the period of grant support, the principal investigator had a number of opportunities to become involved in research in other areas of social and organizational psychology. This additional research reflects the principal investigator's interests in the social influence related to group decision making and the motivational factors that contribute to task performance. A number of projects were completed during the period of support that reflect activities related to these interests which are listed below. These publications and manuscripts involved collaborative efforts with a former student and a colleague who have positions in primarily teaching colleges.

Matz, D.C., & Hinsz, V.B. (2003). Accounting for consistency and change in responses to influence attempts: An examination of preference for consistency. *Current Psychology, 21*, 401-414.

Hinsz, V.B., & Matz, D.C. (2003). Social psychology: A topical review. In *Encyclopedia of Life Support Systems (EOLSS), Psychology*. Oxford, UK: United Nations Educational, Scientific, and Cultural Organization.

Nickell, G.S., Hinsz, V.B., & Park, E.S. (2005). Using normative information to encourage food processing workers to keep food clean. In B. Maunsell & D. J. Bolton (Eds.), *Food Safety Risk Communication: The Message and Motivational Strategies* (pp. 99-109). Dublin, Ireland: Teagasc – The National Food Centre.

Hinsz, V.B., Nickell, G.S., & Park, E.S. (submitted). Work habits' role in the motivation of food safety behaviors. *Organizational Behavior and Human Decision Processes.*

18. **Development of Students**

A number of graduate and undergraduate students were involved in the research efforts described above or whose training was influenced by the research support provided. Below I list the students involved and their current placement. Also, a list of students who completed theses and dissertations is included.

Graduate students:

Dana M. Lawrence – NDSU Psychology graduate student
Jared L. Ladbury – NDSU Psychology graduate student
Melissa Lewis – Post-doctoral Scholar, University of Washington
Sarah Heimerdinger – NDSU Psychology graduate student
Brian Meier – assistant professor, Gettysburg College
Scott Engel – Research Scientist, Neuropsychiatric Research Institute
Brian Gatheridge – graduate student, Washington State University

Undergraduate students:

Casey Spieker – graduate student, Southern Illinois University
Dustin Jundt – graduate student, Michigan State University
Jared Ladbury – graduate student, NDSU
Allison Albrecht – graduate student, Ball State University
Jordan Henkel – pursuing graduate training in public policy
Sarah Heimerdinger – graduate student, NDSU
Brandon Kopp – graduate student, Ohio State University

Among these students, Brian Meier and Dana Lawrence completed master's theses, Dustin Jundt, Casey Spieker, Jared Ladbury, and Sarah Heimerdinger conducted undergraduate honors theses, and Melissa Lewis completed her dissertation during the period of funding.

19. **Presentations Made During Grant Funded Period**

The team of researchers at NDSU involved in this funded project were involved in a large number of presentations during the period of grant support. This generally reflects the principal investigator's encouragement that graduate and undergraduate students gain professional experience by preparing and presenting at the meetings of professional organizations. Although not all of these presentations are directly or indirectly related to the research program being supported, the presentations do indicate research activity that was in part supported by grant funds. Below the list of presentations made during the period of support are listed in chronological order.

Jundt, D. & Hinsz, V.B. (2002, May). Affect influences on mechanisms that mediate the relationship between goals and performance. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Engel, S., & Hinsz, V.B. (2002). Perceptions of men=s preferences in long and short term relationships: What men want and what women think men want. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Hinsz, V.B. (2002). Group decision making and shared task representations. Invited paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Hinsz, V.B. (2002). Research on the groups-as-information-processors perspective. Invited presentation at the Department of Social and Organizational Psychology, University of Amsterdam.

Hinsz, V.B., & Jundt, D.K. (2002). Individual differences in a goal-setting situation: An examination of the Motivational Trait Questionnaire. Paper presented at the meeting of the 25th International Congress of Applied Psychology, Singapore.

Albrecht, A.K., & Hinsz, V.B. (2003). Mock jurors' responsibility judgments and damage awards following the introduction of judgmental anchors. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Heimerdinger, S.R., & Hinsz, V.B. (2003). There are two parts to motivation: The influences of failure-avoidance motivation in a goal-setting situation. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Henkel, J.M., & Hinsz, V.B. (2003). Does success and failure serve as mood induction? Using success and failure in a goal-setting situation to induce mood. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Hinsz, V.B. (2003). Investigation of a model of optimal information integration in groups. Paper presented at the NSF and AFOSR sponsored Workshop on Information Integration in Decision Making, Silver Spring, Maryland.

Spieker, C.J., & Hinsz, V.B. (2003). Repeated success and failure influences on self-efficacy and personal goals. Paper presented at the 15th Annual Convention of the American Psychological Society, Atlanta.

Hinsz, V.B., Meier, B.P., & Gatheridge, B. (2003). Group and individual aggressive actions toward other groups and individuals. Paper presented at the 15th Annual Convention of the American Psychological Society, Atlanta.

Hinsz, V.B. (2003). Group decision making and task performance goals. Invited paper presented at the Group Decision Making: Motivation and Cognition conference, Amsterdam, Netherlands.

Heimerdinger, S.R., & Hinsz, V.B. (2003). An investigation of the tritone paradox with individuals from different linguistic groups. Paper presented at the ND/SD EPSCoR Fourth Biennial Joint Conference, Fargo, North Dakota.

Henkel, J.M., & Hinsz, V.B. (2003). A preliminary investigation of the potential relationships among personality structure, cognitive abilities, and motivational traits. Paper presented at the annual Northern Lights Psychology Conference, Grand Forks, North Dakota.

Heimerdinger, S.R., & Hinsz, V.B. (2003). Affect influences on motivational factors contributing to task performance. Paper presented at the annual Northern Lights Psychology Conference, Grand Forks, North Dakota.

Hinsz, V.B., Heimerdinger, S.R., Henkel, J.M., & Spieker, C.J. (2003). The relevance of base-rate and case-specific information for probability judgments of disease occurrence. Paper presented at the meeting of the Judgment and Decision Making Society, Vancouver, BC.

Hinsz, V.B. (2004). Collaboration in team performance in synthetic task environments. Presentation at the AFOSR sponsored Cognition and Decision Making Workshop, Chandler, AZ.

Hinsz, V.B., & Nickell, G.S. (2004). A motivational model of product safety and security behaviors. Paper presented at the 19th annual meeting of the Society for Industrial and Organizational Psychology, Chicago.

Nickell, G.S., & Hinsz, V.B. (2004). Predicting food safety and security behaviors in turkey processing workers. Paper presented at the 19th annual meeting of the Society for Industrial and Organizational Psychology, Chicago.

Hinsz, V.B., Park, E.S., & Sjomeling, M. (2004). Group interaction sustains positive mood and diminishes negative mood. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Lawrence, D.M., & Hinsz, V.B. (2004). Attributions of responsibility for pooled-performance in groups: Group members take both the actor and observers perspectives. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Spieker, C.J., & Hinsz, V.B. (2004). Two memories are better than one: Collaboration enhances recognition memory performance. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Magnan, R.E., & Hinsz, V.B. (2004). Influence of mood and gender on the level of risk taking advised for others. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Boyce, N.M., & Hinsz, V.B. (2004). The predictive ability of core self-evaluations in a goal-setting situation. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Albrecht, A.K., Hinsz, V.B., & Jundt, D.K. (2004). Predicting individual performance on a sensorimotor task with specific motivational traits. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Henkel, J.M., & Hinsz, V.B. (2004). Satisfaction with performance as a mediator of the relationship between performance outcomes and affective reactions. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Heimerdinger, S.R., Hinsz, V.B., & Jundt, D.K. (2004). Motivational traits and self-consciousness: Exploring the construct validity of the Motivational Trait Questionnaire. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Park, E.S., & Hinsz, V.B. (2004). The influence of coordination and collaboration instructions on UAV team performance. Presentation at the Human Factors in UAVs Workshop, Chandler, AZ.

Heimerdinger, S.R., & Hinsz, V.B. (2004). Exploration of differences in mens' judgments of the likelihood to rape or force sex. Paper presented at the North Dakota Psychological Association, Fargo, ND.

Ladbury, J.L., & Hinsz, V.B. (2004). Social projection in rating the likelihood of desirable and undesirable sexual behaviors. Paper presented at the 16th Annual Convention of the American Psychological Society, Chicago.

Albrecht, A.K., & Hinsz, V.B. (2004). Overestimations of the frequency of common and less-common causes of death. Paper presented at the 16th Annual Convention of the American Psychological Society, Chicago.

Lawrence, D.M., & Hinsz, V.B. (2004). Group member attributions for performance are sensitive to differences in group activities. Paper presented at the 16th Annual Convention of the American Psychological Society, Chicago.

Spieker, C.J., & Hinsz, V.B. (2004). Macro-predictions and test performance for knowledge in physical and mental health domains. Paper presented at the 16th Annual Convention of the American Psychological Society, Chicago.

Heimerdinger, S.R., & Hinsz, V.B. (2004). Dispositional variables influence self-reported likelihood to behave sexually violent toward women. Paper presented at the 16th Annual Convention of the American Psychological Society, Chicago.

Henkel, J.M., & Hinsz, V.B. (2004). The superiority of dyadic memory and metamemory performance. Paper presented at the 16th Annual Convention of the American Psychological Society, Chicago.

Magnan, R.E., & Hinsz, V.B. (2004). Pre and post test metacognitive judgments of mental health knowledge. Paper presented at the 16th Annual Convention of the American Psychological Society, Chicago.

Kaitibie, S., Nganje, W., Hinsz, V., & Nickell, G. (2004). Double moral hazard and food safety risks in turkey processing. Paper presented at the Annual Meeting of the American Agricultural Economics Association, Denver, CO.

Park, E.S., Hinsz, V.B., Magnan, R.E., & Lawrence, D.M. (2004). Affect-specific influences on likelihood judgments of future life-events. Paper presented at the Annual Meeting of the Society for Judgment and Decision Making, Minneapolis, MN.

Magnan, R.E., Hinsz, V.B., Park, E.S., & Lawrence, D.M. (2004). Naturally-occurring moods and the use of judgmental anchors. Paper presented at the Annual Meeting of the Society for Judgment and Decision Making, Minneapolis, MN.

Lawrence, D.M., Hinsz, V.B., Park, E.S., & Magnan, R.E. (2004). Naturally-occurring affective states and framing for gains and losses. Paper presented at the Annual Meeting of the Society for Judgment and Decision Making, Minneapolis, MN.

Spieker, C.J., Hinsz, V.B., & Matz, D.C. (2005). The social and evolutionary psychology of hair color preferences on the Northern and Southern Prairie. Paper presented at the annual meeting of the Society for Personality and Social Psychology, New Orleans, LA.

Hinsz, V.B., Spieker, C.J., Heimerdinger, S.R., & Lawrence, D.M. (2005). The potential of motivation to produce group memory performance beyond optimum. Paper presented at the annual meeting of the Society for Personality and Social Psychology, New Orleans, LA.

Tindale, R.S., Hinsz, V.B., & Henkel, J.M. (2005). Juries and military justice. Paper presented at the Group and Intergroup Processes Preconference of the annual meeting of the Society for Personality and Social Psychology, New Orleans, LA.

Lawrence, D.M., Hinsz, V.B., Park, E.S., Magnan, R.E. (2005). Uncoupling mood and arousal reactions to group interaction. Paper presented at the annual meeting of the Society for Personality and Social Psychology, New Orleans, LA.

Magnan, R.E., Hinsz, V.B., Lawrence, D.M., & Park, E.S. (2005). Trait and state mood in individual and group memory performance. Paper presented at the annual meeting of the Society for Personality and Social Psychology, New Orleans, LA.

Ladbury, J.L., & Hinsz, V.B. (2005). Do individuals have meta-knowledge of group decision processes? Paper presented at the annual meeting of the Society for Personality and Social Psychology, New Orleans, LA.

Park, E.S., Hinsz, V.B., & Nickell, G.S. (2005). Exploring the differences between promotion and prevention focused employees. Paper presented at the 20th annual meeting of the Society for Industrial and Organizational Psychology, Los Angeles.

Morgan, J., Nickell, G., & Hinsz, V. (2005). Employees' willingness to prevent intentional food contamination. Poster presented at the Red River Psychology Conference, Moorhead, Minnesota.

Hinsz, V.B. (2005). Modeling team performance in a UAV ground control station task environment. Paper presented at the Air Force Office of Scientific Research Workshop on Cognition and Decision, St. Augustine, FL.

Henkel, J.M., & Hinsz, V.B. (2005). Males were disproportionately elected as jury forepersons in mock trials. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.

Heimerdinger, S.R., & Hinsz, V.B. (2005). Psychometric properties of the Stress in General Scale. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.

Park, E.S., Hinsz, V.B. (2005). Do groups accentuate mood-congruent memory effects? Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.

Ladbury, J.L, & Hinsz, V.B. (2005). Using cultural consensus to test cultural differences of perceived hair color preferences. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.

Hinsz, V.B. (2005). BRUTE: A versatile research platform for studying UAV operator behavior. Invited paper presented at the 2nd annual Human Factors of UAVs Workshop, Mesa, AZ.

Heimerdinger, S.R., & Hinsz, V.B. (2005). Peer discussion reduces self-reported intentions to participate in group sexual assault. Presented at the annual meeting of the American Psychological Society, Los Angeles.

Logue, C.M, Nolan, L.K., Hinsz, V.B., Nganje, W., Sellnow, T., & Khaitsa, M. (2005). Food safety risk assessment at NDSU – Finding links. Presented at USDA CSREES Post-Award Management Workshop for Non-Competitive Food Safety Grants, Washington DC.

Hinsz, V.B. (2005). Prospects for theory, method, and research on affect in groups. Presented in the Getting in the Mood with Groups: Integrating Affect and Groups symposium at the annual meeting of the Society for Experimental Social Psychology, San Diego.

Hinsz, V.B. (2005). Commitment in group decision making. Presented at the conference on Commitment in Organizations: Accumulated Wisdom and New Directions, Columbus, OH.

SYMPOSIA and PRECONFERENCE ORGANIZING

Group and Intergroup Processes Preconference of the 2005 annual meeting of the Society for Personality and Social Psychology, New Orleans, LA.

Strategies of Information Processing in Groups symposium at the 2005 anual meeting of the Midwestern Psychological Association, Chicago, IL.

Getting in the Mood with Groups: Integrating Affect and Groups symposium at the 2005 annual meeting of the Society for Experimental Social Psychology, San Diego.

Metacognition in Groups: How Metacognitive Beliefs Facilitate Group Interaction and Performance at the 2006 annual meeting of the Society for Personality and Social Psychology, Palm Springs, CA.

20. **Literature Cited**

Asch, S.E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied, 70*, 1-70.

Baddeley, A.D. (1986). *Working memory.* Oxford: Oxford University Press.

Baddeley, A.D. (1992). Working memory. *Science, 255*, 556-559.

Banks, W.P. (1970). Signal detection theory and human memory. *Psychological Bulletin, 74*, 81-99.

Basden, B.H., Basden, D.R., & Henry, S. (2000). Costs and benefits of collaborative remembering. *Applied Cognitive Psychology, 14*, 497-507.

Bjork, R.A. (2001). Recency and recovery in human memory. In H.L. Roediger, III, J.S. Nairne, I. Neath, & A.M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 221-232). Washington, DC: American Psychological Association.

Burns, D.J., & Gold, D.E. (1999). An analysis of item gains and losses in retroactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 978-985.

Clark, N.K., & Stephenson, G.M. (1989). Group remembering. In P. Paulus (Ed.), *Psychology of group influence* (2nd. ed., pp. 357-391). Hillsdale, NJ: Erlbaum.

Clark, S.E., Hori, A., Putnam, A., & Martin, T.P. (2000). Group collaboration in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1578-1588.

Collins, B.E., & Guetzkow, H. (1964). *A social psychology of group processes for decision-making.* New York: John Wiley & Sons.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Davis, J.H. (1969). *Group performance.* Reading, MA: Addison-Wesley.

Davis, J.H. (1973). Group decision and social interaction: A theory of social decision schemes. *Psychological Review, 80*, 97-125.

Davis, J.H. (1980). Group decision and procedural justice. In M. Fishbein (Ed.), *Progress in social psychology* (pp. 157-229). Hillsdale, NJ: Erlbaum.

Davis, J.H. (1982). Social interaction as a combinatorial process in group decision. In H. Brandstatter, J.H. Davis, & G. Stocker-Kreichgauer (Eds.), *Group decision making* (pp. 27-58). London: Academic Press.

Davis, J.H. (1992). Some compelling intuitions about group consensus decisions, theoretical and empirical research, and interpersonal aggregation phenomena: Selected examples, 1950-1990. *Organizational Behavior and Human Decision Processes, 52*, 3-38.

Efron, R. (1970). The minimum duration of a perception. *Neuropsychologia, 8*, 37-55.

Entin, E.B., Entin, E.E., Bailey, A., Paley, M., & Miller, D. (2001). *A system to enhance team decision making performance.* Technical Report. Woburn, MA: Aptima, Inc.

Entin,. E.E., & Rubineau, B. (2002). *AWACS dynamic distributed decision making synthetic task environment.* Technical Report. Woburn, MA: Aptima, Inc.

Fahey, R.P., Rowe, A.L., Dunlap, K.L., & deBoom, D.O. (2000). *Synthetic Task Design (1): Preliminary cognitive task analysis of AWACS weapons director teams.* Report submitted to the Air Force Research Laboratory, Human Effectiveness Directorate, Brooks Air Force Base, TX.

Green, D.M., & Swets, J.A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.
Guzzo, R.A., & Salas, E. (1995). *Team effectiveness and decision making in organizations.* San Francisco: Jossey-Bass.
Haber, R.N., & Nathanson, L.S. (1968). Post-retinal storage? Park's camel as seen through the eye of a needle. *Perception and Psycholophysics, 3,* 349-355.
Hackman, J.R. (1987). The design of work teams. In J.W. Lorsch (Ed.), *Handbook of organizational behavior* (pp. 315-342). Englewood Cliffs, NJ: Prentice-Hall.
Hackman, J.R. (1990). *Groups that work (and those that don't).* San Francisco: Jossey-Bass.
Hackman, J.R. (1992). Group influences on individuals in organizations. In M.D. Dunnette & L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology (2nd ed., Vol. 3, pp. 199-267).* Palo Alto, CA: Consulting Psychologists Press.
Hartwick, J., Sheppard, B.H., & Davis, J.H. (1982). Group remembering: Research and implications. In R.A. Guzzo (Ed.), *Improving group decision making in organizations* (pp. 41-72). New York: Academic Press.
Hastie, R. (1986). Experimental evidence on group accuracy. In B. Grofman & G. Owen (Eds.) *Decision Research* (Vol. 2, pp. 129-157). Greenwich, CT: JAI Press.
Heggestad, E.D., & Kanfer, R. (2000). Individual differences in trait motivation: Development of the Motivational Trait Questionnaire. *International Journal of Educational Research, 33,* 751-776.
Hertel, G., Kerr, N.L., & Messé, L.A. (2000). Motivation gains in performance groups. Paradigmatic and theoretical developments on the Köhler Effect. *Journal of Personality and Social Psychology, 79,* 580-601.
Hill, G.W. (1982). Group versus individual performance: Are N +1 heads better than one? *Psychological Bulletin, 91,* 517-539.
Hinsz, V. B. (1990). Cognitive and consensus processes in group recognition memory performance. *Journal of Personality and Social Psychology, 59,* 705-718.
Hinsz, V.B. (1999). Group decision making with responses of a quantitative nature: The theory of social decision schemes for quantities. *Organizational Behavior and Human Decision Processes, 80,* 28-49.
Hinsz, V.B., Heimerdinger, S.R., Henkel, J.M., & Spieker, C.J. (2005). Test-accuracy and base-rate information in the prediction of disease occurrence. *Current Psychology, 24,* 80-90.
Hinsz, V.B., & Malone, C.P. (2004). Memory for attributes of information presented in a synthetic task environment: An illustration with AWACS weapons directors' displays. In S.G. Schiflett, L.R., Elliott, E. Salas, & M.D. Coovert (Eds.), *Scaled Worlds: Development, validation, and applications* (pp. 220-239). Hampshire, England: Ashgate Publishing Limited.
Hinsz, V.B., Spieker, C.J., Heimerdinger, S.R., & Lawrence, D.M. (2005). The potential of motivation to produce group memory performance beyond optimum. Paper presented at the annual meeting of the Society for Personality and Social Psychology, New Orleans, LA.
Hinsz, V.B., Tindale, R.S., & Vollrath, D.A. (1997) The emerging conceptualization of groups as information processors. *Psychological Bulletin, 121,* 43-64.
Hinsz, V. B., Vollrath, D. A., Nagao, D. H. & Davis, J. H. (1988). Comparing the structure of individual and small group perceptions. *International Journal of Small Group Research,* 4, 159-168.

Hollingshead, A.B. (1998). Retrieval processes in transactive memory systems. *Journal of Personality and Social Psychology, 74*, 659-671.

Ilgen, D.R. (1999). Teams embedded in organizations. *American Psychologist, 54*, 129-139.

International Personality Item Pool (2001). A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences (http://ipip.ori.org/). Internet Web Site.

Karau, S.J., & Williams, K.D. (1997). The effects of group cohesiveness on social loafing and social compensation. *Group Dynamics, 1,* 156-168.

Kerr, N.L., Stasser, G., Davis, J.H. (1979). Model testing, model fitting, and social decision schemes. *Organizational Behavior and Human Performance, 23*, 399-410.

Kirchler, E., & Davis, J.H. (1986). The influence of member status differences and task type on group consensus and member position change. *Journal of Personality and Social Psychology, 51*, 83-91.

Kogan, N., & Wallach, M.A. (1964). *Risk-taking: A Study in Cognition and Personality.* New York: Holt, Rhinehart and Winston.

Latané, B., Williams, K.D., & Harkins, S.G. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology, 37*, 822-832.

Laughlin, P.R. (1980). Social combination processes in cooperative problem-solving groups on verbal intellective tasks. In M. Fishbein (Ed.), *Progress in social psychology* (pp. 127-155). Hillsdale, NJ: Erlbaum.

Laughlin, P.R. (1996). Group decision making and collective induction. In E. Witte & J.H. Davis (Eds.), *Understanding group behavior: Consensual action by small groups,* (Vol. 1., pp. 61-80). Hillsdale, NJ: Erlbaum.

Locke, E., Feren, D., McCaleb, V., Shaw, K., & Denny, A. (1980). The relative effectiveness of four methods of motivating employee performance. In K. Duncan, M. Gruenberg, & D. Wallis (Eds.), *Changes in working life,* pp.363-388. New York: Wiley.

Lorge, I., & Solomon, H. (1955). Two models of group behavior in the solution of Eureka-type problems. *Psychometrica, 20*, 139-148.

Madigan, S. (1983). Picture memory. In J.C. Yuille (Ed.), *Imagery, memory, and cognition: Essays in honour of Allan Paivio* (pp. 65-89). Hillsdale, NJ: Erlbaum.

McGrath, J.E. (1984). *Groups: Interaction and performance.* Englewood Cliffs, NJ: Prentice-Hall.

Meudell, P.R., Hitch, G.J., & Bolye, M.M. (1995). Collaboration in recall: Do pairs of people cross-cue each other to produce new memories? *Quarterly Journal of Experimental Psychology, 48A*, 141-152.

Miller, G.A (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Moreland, R.L. (1999). Transactive memory: Learning who knows what in work groups and organizations. In L. Thompson, D. Messick, & J. Levine (Eds.), *Shared cognition in organizations: The management of knowledge* (pp. 3-31). Mahwah, NJ: Erlbaum.

Paulus, P. B., Dugosh, K. L., Dzindolet, M. T., Coskun, H., & Putman, V. L. (2002). Social and cognitive influences in group brainstorming. Predicting production gains and losses. *European Review of Social Psychology, 12*, 299-325.

Peterson, L.R., & Peterson, M.J. (1959). Short-term retention of individual items. *Journal of Experimental Psychology, 61*, 12-21.

Schacter, D.L., & Tulving, E. (1994). What are the memory systems of 1994? In D.L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 1-38). Cambridge, MA: MIT Press.

Shaw, Marvin E. 1981. Group dynamics: The psychology of small group behavior (3rd ed.). New York: McGraw-Hill.

Shepperd, J. A. (1993). Productivity loss in groups: A motivation analysis. *Psychological Bulletin, 113*, 67-81.

Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.

Sorkin, R.D., & Dai, H. (1994). Signal detection analysis of the ideal group. *Organizational Behavior and Human Decision Processes, 60,* 1-13.

Sorkin, R.D., Hays, C.J., & West, R. (2001). Signal-detection analysis of group decision making. *Psychological Review, 108*, 183-203.

Sorkin, R.D., West, R., & Robinson, D.E. (1998). Group performance depends on the majority rule. *Psychological Science, 9*, 456-463.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74,* 1-29.

Stasser, G. (1999). A primer of social decision scheme theory: Models of group influence, competitive model testing, and prospective modeling. *Organizational Behavior and Human Decision Processes, 80*, 3-20.

Steiner, I.D. (1972). *Group process and productivity*. New York: Academic Press.

Teasdale, J.D., Dritschel, B.H., Taylor, M.J., Proctor, L., et al. (1995). Stimulus-independent thought depends on central executive resources. *Memory & Cognition, 23*, 551-559.

Tetlock, P. E. (1985). Accountability: The neglected social context of judgment and choice. In B. Staw & L. Cummings (Eds.), *Research in organizational behavior* (Vol. 7, pp. 297-332). Greenwich, CT: JAI Press.

Tulving, E. (1974). Cue-dependent forgetting. *American Scientist, 62*, 74-82.

Vollrath, D. A., Sheppard, B. H., Hinsz, V. B. & Davis, J. H. (1989). Memory performance by decision-making groups and individuals. *Organizational Behavior and Human Decision Processes, 43*, 289-300.

Wallach, M. A., Kogan, N., & Bem, D. J. (1964). Diffusion of responsibility and level of risk taking in groups. *Journal of Abnormal and Social Psychology, 68*, 263-274.

Wegner, D.M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G.R. Goethals (Eds.), *Theories of group behavior* (pp. 185-208). New York: Springer-Verlag.

Wegner, D.M. (1995). A computer network model of human transactive memory. *Social Cognition, 13*, 319-339.

Weldon, M.S., Blair, C., & Huebsch, P.D. (2000). Group remembering: Does social loafing unlie collaborative inhibition? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1568-1577.

21. **Personnel Supported by Grant Funds**

Below are listed the individuals who were directly supported by grant funds during some portion of the funded period.

Verlin B. Hinsz, Principal Investigator.
Ryan Pillatzki, technician and programmer
Ernest S. Park, post-doctoral scholar
Christine Malone, senior research associate
Magdalene Chalikia, senior research associate
Wendy Troop-Gordon, senior research associate
Dana Lawrence, graduate research assistant
Jordan Henkel, undergraduate research assistant
Casey Spieker, undergraduate research assistant
Leyna Deal, undergraduate research assistant
Allison Albrecht, undergraduate research assistant
Alice Scharnweber, undergraduate research assistant
Mary Blume, undergraduate research assistant
Marina Serdiouk, undergraduate research assistant

22. **Appendix A**

Image 1. Laboratory set up showing one participant sitting at the computer workstation.

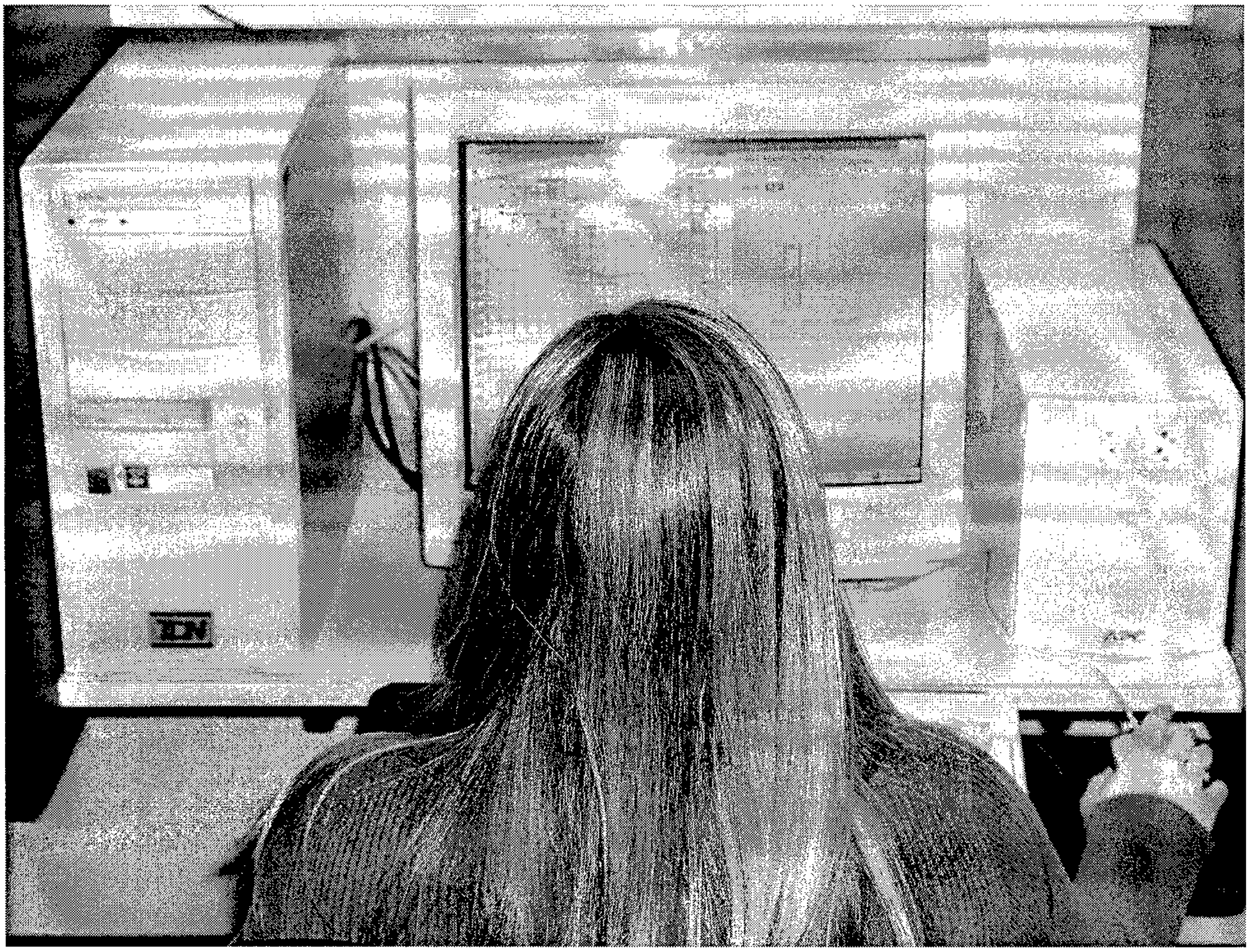


Image 2. Laboratory set up showing five participants sitting at the computer workstations for the five-person teams.

Image 3. Laboratory set up showing three participants sitting at the computer workstations for the three-person teams.

Image 4. The workstations that control the operations of the computer workstations assigned to the three- and five-person teams.

23. Appendix B

Image 5. The layout of the screen display for participants at the beginning of the scenario. The screen display provides the controls for the weapons directors' activities as well as the grid reflecting the locations of the aircraft and defined locations on the map.

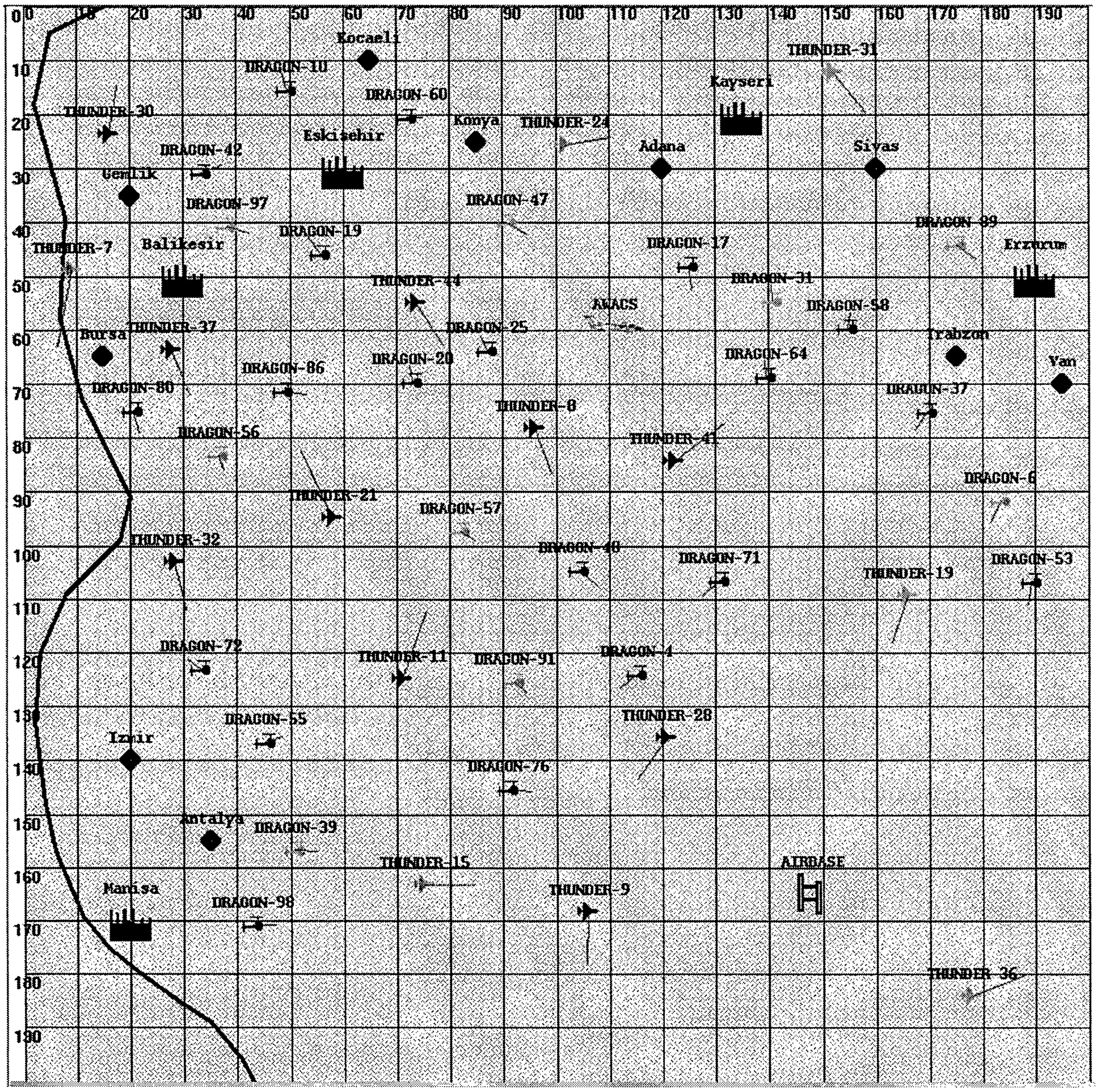


Image 6. A sample screen display of the participants serving as weapons directors.

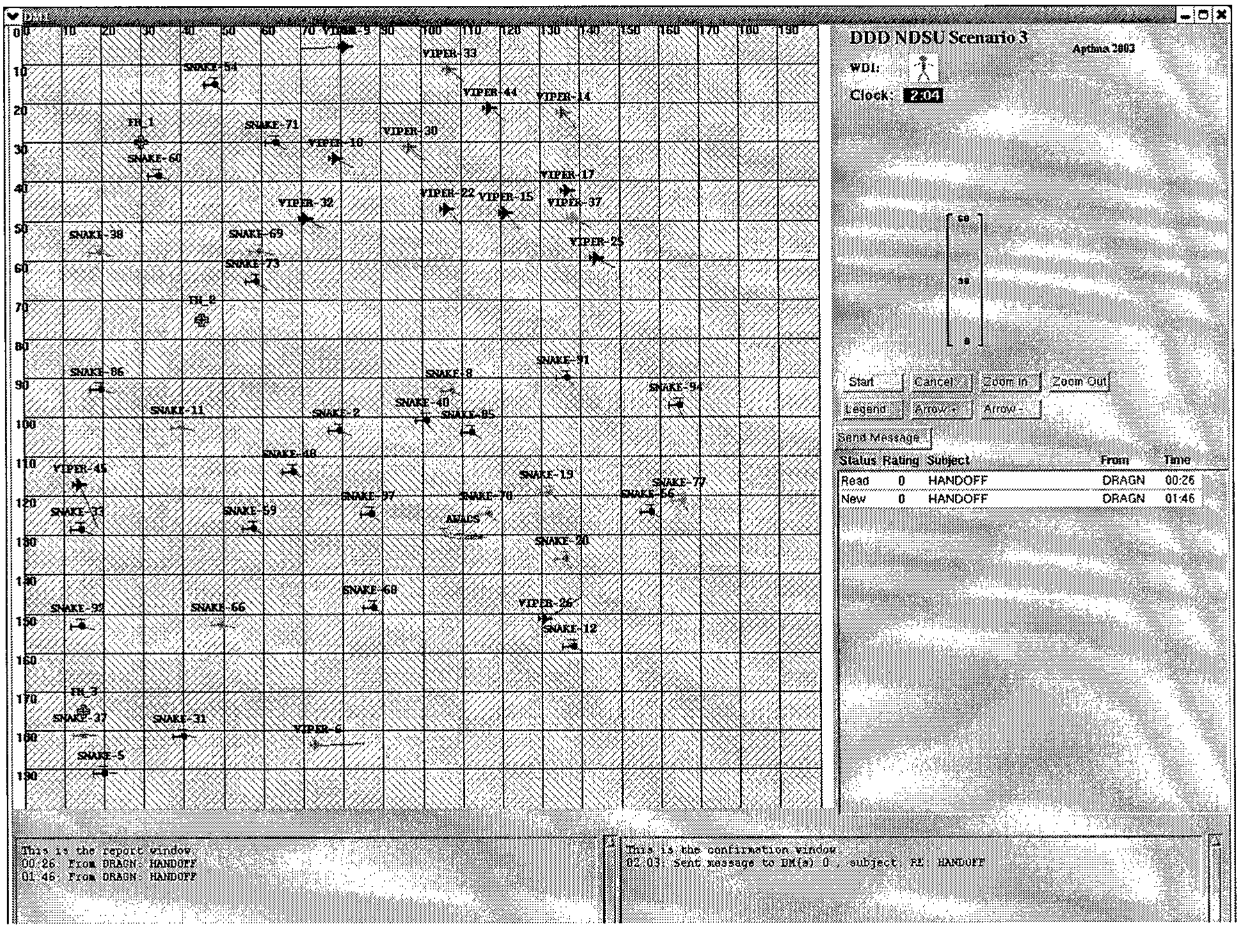


Image 7. The pop-up menu of information about an aircraft.

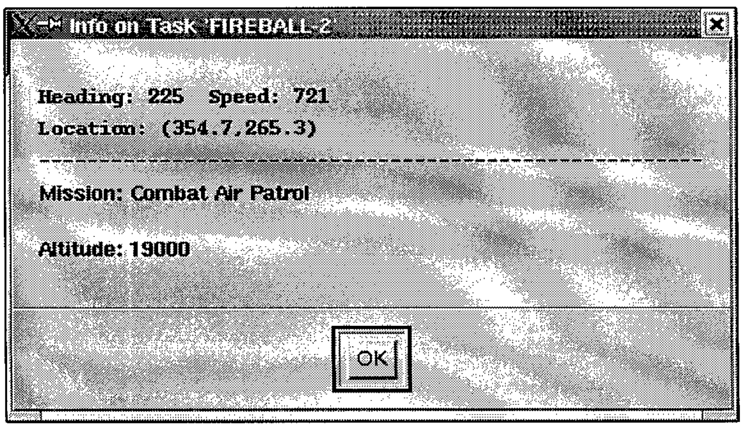
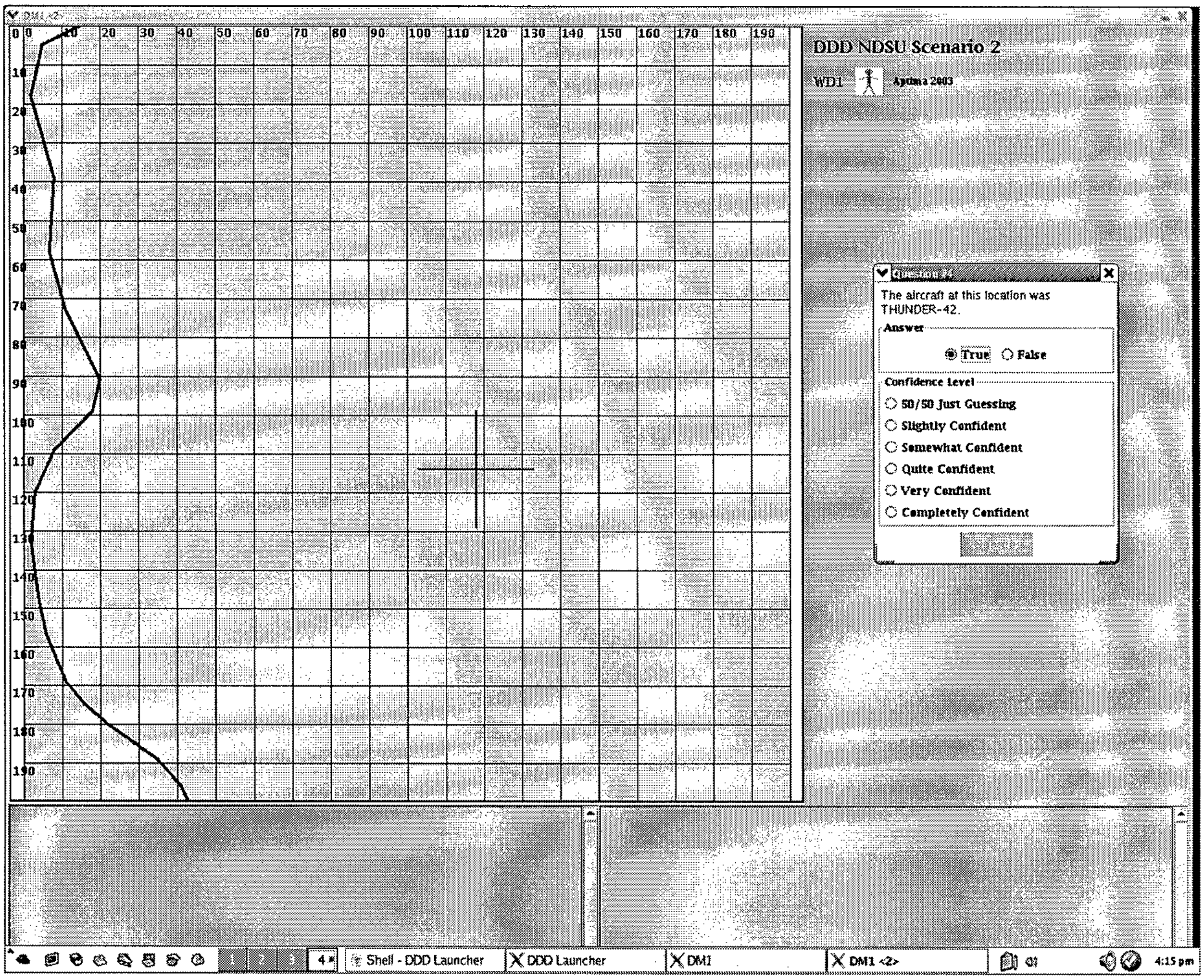


Image 8. The screen display of the map grid that is used to gather the recognition item responses.



Coleman, C.A., B.E. Hull, J.N. McDougal and J.V. Rogers. The effect of m-xylene on cytotoxicity and cellular antioxidant status in rat dermal equivalents. *Toxicology Letters*, 142:133-142 (2003).

Gunasekar, P.G., J.V. Rogers, M.H. Kabbur, C.M. Garrett, W.W. Brinkley and J.N. McDougal. Molecular and Histological Responses Related to Acute Skin Irritation from Exposure to m-Xylene. *Journal Biochemical and Molecular Toxicology,* 17(2):92-94 (2003).

Rogers, J.V., C.M. Garrett and J.N. McDougal. Gene Expression in Rat Skin Induced by Irritating Chemicals. *Journal Biochemical and Molecular Toxicology*, 17(3):123-137 (2003).

Rogers, J.V., C.M. Garrett and J.N. McDougal. Gene Expression in Rat Skin Induced by Irritating Chemicals. *Journal Biochemical and Molecular Toxicology*, 17(3):123-137 (2003).

McDougal, J.N.: Physiologically-based Pharmacokinetic Modeling, in Dermatotoxicology, F.N. Marzulli and H.I. Maibach (EDS), 6th edition, Hemisphere Publishing Corporation, Washington DC, (2004).

McDougal, J.N. and J.V. Rogers. Local and Systemic Toxicity of JP-8 from Cutaneous Exposures. *Toxicology Letters.* 149(1-3):301-308 (2004).

Rogers, J.V., G.L. Seigel, D.L. Pollard, A.D. Rooney and J.N. McDougal. The Cytotoxicity of JP-8 volatile jet fuel components in keratinocytes. *Toxicology.* 197(2):113-121 (2004).

Rogers, J.V., G.L. Seigel, D.L. Pollard, A.D. Rooney and J.N. McDougal. The Cytotoxicity of JP-8 volatile jet fuel components in keratinocytes. *Toxicology.* 197(2):113-121 (2004).

McDougal, J.N.: Physiologically-based Pharmacokinetic Modeling, in Dermatotoxicology, H. Zhai and H.I. Maibach (EDS), 6th edition, CRC Press, Boca Raton FL, 591-619 (2004).

McDougal, J.N. and J.V. Rogers. Local and Systemic Toxicity of JP-8 from Cutaneous Exposures. *Toxicology Letters.* 149(1-3):301-308 (2004).

McDougal, J.N. C.M. Amato, and C.M. Garrett. Molecular Changes in Rat Skin related to Irritation by JP-8 Jet Fuel. The Toxicologist 78 (S1) p. 325 (2004).

McDougal, J.N. Zheng, Y, Zhang, Q. and Conolly, R. Biologically Based Pharmacokinetic and Pharmacodynamic Models of the Skin, in Dermal Absorption Models in Toxicology and Pharmacology, J.E. Riviere (ed), Taylor and Francis, London UK (2005).

Inventions/Patents/Discoveries:

none

Collaborators/Consultants:

Rory Conolly, CIIT Centers for health research (now USEPA)
Jeff Fisher, University of Georgia
James Rogers, Battelle Columbus
Mandip Sachdeva, Florida A&M University
Frank Witzmann, Indiana University School of Medicine